

## INTERNATIONAL SEARCH REPORT

(PCT Article 18 and Rules 43 and 44)

|   |   |   |
|---|---|---|
| Applicant's or agent's file reference<br>FB/BC45215     | <b>FOR FURTHER ACTION</b> see Notification of Transmittal of International Search Report (Form PCT/ISA/220) as well as, where applicable, item 5 below. |   |
| International application No.<br>PCT/EP 00/ 00346       | International filing date (day/month/year)<br>17/01/2000  | (Earliest) Priority Date (day/month/year)<br>19/01/1999 |
| Applicant<br>SMITHKLINE BEECHAM BIOLOGICALS S.A. et al. |   |   |

This International Search Report has been prepared by this International Searching Authority and is transmitted to the applicant according to Article 18. A copy is being transmitted to the International Bureau.

This International Search Report consists of a total of 5 sheets.

☐ It is also accompanied by a copy of each prior art document cited in this report.

## 1. Basis of the report

- a. With regard to the language, the international search was carried out on the basis of the international application in the language in which it was filed, unless otherwise indicated under this item.

☐ the international search was carried out on the basis of a translation of the international application furnished to this Authority (Rule 23.1(b)).

- b. With regard to any nucleotide and/or amino acid sequence disclosed in the international application, the international search was carried out on the basis of the sequence listing :

☐ contained in the international application in written form.

☐ filed together with the international application in computer readable form.

☒ furnished subsequently to this Authority in written form.

☒ furnished subsequently to this Authority in computer readable form.

☒ the statement that the subsequently furnished written sequence listing does not go beyond the disclosure in the international application as filed has been furnished.

☒ the statement that the information recorded in computer readable form is identical to the written sequence listing has been furnished

2. ☐ Certain claims were found unsearchable (See Box I).

3. ☒ Unity of invention is lacking (see Box II).

## 4. With regard to the title,

☐ the text is approved as submitted by the applicant.

☒ the text has been established by this Authority to read as follows:

POLYNUCLEOTIDES AND THEIR APPLICATION IN DIAGNOSTICS

## 5. With regard to the abstract,

☒ the text is approved as submitted by the applicant.

☐ the text has been established, according to Rule 38.2(b), by this Authority as it appears in Box III. The applicant may, within one month from the date of mailing of this international search report, submit comments to this Authority.

## 6. The figure of the drawings to be published with the abstract is Figure No.

☐ as suggested by the applicant.

☐ because the applicant failed to suggest a figure.

☐ because this figure better characterizes the invention.

☒ None of the figures.

# INTERNATIONAL SEARCH REPORT

International application No.  
PCT/EP 00/00346

## Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This International Search Report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:  
because they relate to subject matter not required to be searched by this Authority, namely:
2. ☐ Claims Nos.:  
because they relate to parts of the International Application that do not comply with the prescribed requirements to such an extent that no meaningful International Search can be carried out, specifically:
3. ☐ Claims Nos.:  
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

## Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

*see additional sheet*

1. ☐ As all required additional search fees were timely paid by the applicant, this International Search Report covers all searchable claims.
2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.
3. ☐ As only some of the required additional search fees were timely paid by the applicant, this International Search Report covers only those claims for which fees were paid, specifically claims Nos.:
4. ☒ No required additional search fees were timely paid by the applicant. Consequently, this International Search Report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:  
  
1-5 all partially

Remark on Protest

- ☐ The additional search fees were accompanied by the applicant's protest.  
☐ No protest accompanied the payment of additional search fees.

## FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

## 1. Claims: 1-5 all partially

A polynucleotide comprising a nucleotide sequence which has at least 70% identity to that of Seq.ID:1, a nucleotide sequence complementary to said polynucleotide, or fragments of said polynucleotide. A polynucleotide as above, in which the identity to Seq.ID:1 is at least 95%. A polynucleotide selected from: the polynucleotide of Seq.ID:1, a polynucleotide obtainable by screening an appropriate library with a labeled probe having the sequence of Seq.ID:1 or a fragment thereof, fragments or complementary nucleotide sequences thereof. A process for diagnosing a disease or susceptibility of disease, namely colon cancer, related to the presence or expression or activity of a polynucleotide, comprising analyzing the presence or amount of said polynucleotide in a sample derived from a subject.

## 2. Claims: 1-5 all partially

As invention 1 but concerning Seq.ID:2

## 3. Claims: 1-5 all partially

As invention 1 but concerning Seq.ID:3

## 4. Claims: 1-5 all partially

As invention 1 but concerning Seq.ID:4

## 5. Claims: 1-5 all partially

As invention 1 but concerning Seq.ID:5

## 6. Claims: 1-5 all partially

As invention 1 but concerning Seq.ID:6

## INTERNATIONAL SEARCH REPORT

International Application No

PCT/EP 00/00346

## A. CLASSIFICATION OF SUBJECT MATTER

IPC 7 C12N15/11 C12N15/12 C12Q1/68

According to International Patent Classification (IPC) or to both national classification and IPC

## B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC 7 C12N C12Q

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

## C. DOCUMENTS CONSIDERED TO BE RELEVANT

| Category * | Citation of document, with indication, where appropriate, of the relevant passages  | Relevant to claim No. |
|------------|---|-----------------------|
| X          | Nucleotide Sequence Database EMBL<br>ID HS1220328, Accession number AA422178<br>19 may 1997<br>99.8% identity with Seq.ID:1 nt.281-784<br>reverse orientation<br>XP002136186<br>the whole document<br>--- | 1,3                   |
| X          | Nucleotide Sequence Database EMBL<br>ID AA587764, Accession number AA587764<br>18 september 1997<br>99.3% identity with Seq.ID:1 nt.1-441<br>XP002136187<br>the whole document<br>---<br>-/--             | 1,3                   |



Further documents are listed in the continuation of box C.



Patent family members are listed in annex.

## \* Special categories of cited documents :

- \*A\* document defining the general state of the art which is not considered to be of particular relevance
- \*E\* earlier document but published on or after the international filing date
- \*L\* document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
- \*O\* document referring to an oral disclosure, use, exhibition or other means
- \*P\* document published prior to the international filing date but later than the priority date claimed

- \*T\* later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
- \*X\* document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
- \*Y\* document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.
- \*&\* document member of the same patent family

Date of the actual completion of the international search

25 April 2000

Date of mailing of the international search report

13. 07. 00

Name and mailing address of the ISA

European Patent Office, P.B. 5818 Patentlaan 2  
NL - 2280 HV Rijswijk  
Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,  
Fax: (+31-70) 340-3016

Authorized officer

Macchia, G

## INTERNATIONAL SEARCH REPORT

International Application No

PCT/EP 00/00346

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

| Category * | Citation of document, with indication, where appropriate, of the relevant passages   | Relevant to claim No. |
|------------|--|-----------------------|
| X          | Nucleotide Sequence Database EMBL<br>ID AI339648, Accession number AI339648<br>31 december 1998<br>100% identity with Seq.ID:1 nt.6-411<br>XP002136188<br>the whole document | 1,3                   |
| A          | ---<br>WO 98 53319 A (THE JOHNS HOPKINS<br>UNIVERSITY (US); VOGELSTEIN BERT; KINZLER<br>KENNETH W.) 26 November 1998 (1998-11-26)<br>abstract<br>page 22 -page 80            | 4,5                   |
| A          | ---<br>WO 97 33903 A (SMITHKLINE BEECHAM CORP<br>(US); HUMAN GENOME SCIENCES INC (US); WEI<br>Y-F) 18 September 1997 (1997-09-18)<br>abstract<br>page 50; claim 19           | 4,5                   |
| A          | ---<br>WO 95 15381 A (JOHNS HOPKINS UNIV (US); DE<br>LA CHAPELLE A.; VOGELSTEIN B.; KINZLER<br>K.W) 8 June 1995 (1995-06-08)<br>abstract<br>page 48; claim 11<br>-----       | 4,5                   |

# INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No

PCT/EP 00/00346

| Patent document<br>cited in search report |   | Publication<br>date | Patent family<br>member(s) |           | Publication<br>date |
|---|---|---------------------|----------------------------|-----------|---------------------|
| WO 9853319                                | A | 26-11-1998          | AU                         | 7499198 A | 11-12-1998          |
| WO 9733903                                | A | 18-09-1997          | AU                         | 5305596 A | 01-10-1997          |
| WO 9515381                                | A | 08-06-1995          | EP                         | 0730648 A | 11-09-1996          |
|   |   |                     | JP                         | 9506509 T | 30-06-1997          |
|   |   |                     | US                         | 5693470 A | 02-12-1997          |
|   |   |                     | US                         | 5837443 A | 17-11-1998          |

## PATENT COOPERATION TREATY

PCT

## NOTIFICATION OF ELECTION

(PCT Rule 61.2)

From the INTERNATIONAL BUREAU

To:

Assistant Commissioner for Patents  
 United States Patent and Trademark  
 Office  
 Box PCT  
 Washington, D.C.20231  
 ETATS-UNIS D'AMERIQUE

in its capacity as elected Office

|   |   |
|---|---|
| <b>Date of mailing</b> (day/month/year)<br>11 September 2000 (11.09.00)         |   |
| <b>International application No.</b><br>PCT/EP00/00346                          | <b>Applicant's or agent's file reference</b><br>FB/BC45215          |
| <b>International filing date</b> (day/month/year)<br>17 January 2000 (17.01.00) | <b>Priority date</b> (day/month/year)<br>19 January 1999 (19.01.99) |
| <b>Applicant</b><br>VINALS-BASSOLS, Carlota                                     |   |

1. The designated Office is hereby notified of its election made:



in the demand filed with the International Preliminary Examining Authority on:

26 July 2000 (26.07.00)



in a notice effecting later election filed with the International Bureau on:

2. The election ☒ was

was not

made before the expiration of 19 months from the priority date or, where Rule 32 applies, within the time limit under Rule 32.2(b).

The International Bureau of WIPO  
 34, chemin des Colombettes  
 1211 Geneva 20, Switzerland

Facsimile No.: (41-22) 740.14.35

Authorized officer

C. Cupello

Telephone No.: (41-22) 338.83.38

# PATENT COOPERATION TREATY

From the  
INTERNATIONAL PRELIMINARY EXAMINING AUTHORITY

## PCT

To:  
  
Tyrell, Arthur W R.  
SMITHKLINE BEECHAM  
Two New Horizons Court  
Brentford  
Middlesex TW8 9EP  
GRANDE BRETAGNE

RECEIVED  
  
18 JUN 2001  
  
NEW HORIZONS COURT

NOTIFICATION OF TRANSMITTAL OF  
THE INTERNATIONAL PRELIMINARY  
EXAMINATION REPORT  
(PCT Rule 71.1)

**Fax no: 44 20 8975 6294**

|                                     |            |
|-------------------------------------|------------|
| Date of mailing<br>(day/month/year) | 15.06.2001 |
|-------------------------------------|------------|

|   |                               |
|---|-------------------------------|
| Applicant's or agent's file reference<br>FB/BC45215 | <b>IMPORTANT NOTIFICATION</b> |
|---|-------------------------------|

|   |  |  |
|---|--|--|
| International application No.<br>PCT/EP00/00346 | International filing date (day/month/year)<br>17/01/2000 | Priority date (day/month/year)<br>19/01/1999 |
|---|--|--|

|   |
|---|
| Applicant<br>SMITHKLINE BEECHAM BIOLOGICALS S.A. et al. |
|---|


1. The applicant is hereby notified that this International Preliminary Examining Authority transmits herewith the international preliminary examination report and its annexes, if any, established on the international application.
2. A copy of the report and its annexes, if any, is being transmitted to the International Bureau for communication to all the elected Offices.
3. Where required by any of the elected Offices, the International Bureau will prepare an English translation of the report (but not of any annexes) and will transmit such translation to those Offices.

#### 4. REMINDER

The applicant must enter the national phase before each elected Office by performing certain acts (filing translations and paying national fees) within 30 months from the priority date (or later in some Offices) (Article 39(1)) (see also the reminder sent by the International Bureau with Form PCT/IB/301).

Where a translation of the international application must be furnished to an elected Office, that translation must contain a translation of any annexes to the international preliminary examination report. It is the applicant's responsibility to prepare and furnish such translation directly to each elected Office concerned.

For further details on the applicable time limits and requirements of the elected Offices, see Volume II of the PCT Applicant's Guide.

|  |  |
|--|--|
| <p>Name and mailing address of the IPEA/</p> <p> European Patent Office - P.B. 5818 Patentlaan 2<br/>NL-2280 HV Rijswijk - Pays Bas<br/>Tel. +31 70 340 - 2040 Tx: 31 651 epo nl<br/>Fax: +31 70 340 - 3016</p> | <p>Authorized officer</p> <p>Cardenas, C</p> <p>Tel. +31 70 340-3370</p> |
|--|--|





# PATENT COOPERATION TREATY

## PCT

### INTERNATIONAL PRELIMINARY EXAMINATION REPORT

(PCT Article 36 and Rule 70)

|  |  |  |   |
|--|--|--|---|
| Applicant's or agent's file reference<br><b>FB/BC45215</b>                                       | <b>FOR FURTHER ACTION</b>  |  | See Notification of Transmittal of International Preliminary Examination Report (Form PCT/IPEA/416) |
| International application No.<br><b>PCT/EP00/00346</b>   | International filing date ( <i>day/month/year</i> )<br><b>17/01/2000</b> | Priority date ( <i>day/month/year</i> )<br><b>19/01/1999</b> |   |
| International Patent Classification (IPC) or national classification and IPC<br><b>C12N15/11</b> |  |  |   |
| Applicant<br><b>SMITHKLINE BEECHAM BIOLOGICALS S.A. et al.</b>                                   |  |  |   |

1. This international preliminary examination report has been prepared by this International Preliminary Examining Authority and is transmitted to the applicant according to Article 36.


2. This REPORT consists of a total of 9 sheets, including this cover sheet.

☒ This report is also accompanied by ANNEXES, i.e. sheets of the description, claims and/or drawings which have been amended and are the basis for this report and/or sheets containing rectifications made before this Authority (see Rule 70.16 and Section 607 of the Administrative Instructions under the PCT).

These annexes consist of a total of 1 sheets.

3. This report contains indications relating to the following items:

- I ☒ Basis of the report
- II ☐ Priority
- III ☒ Non-establishment of opinion with regard to novelty, inventive step and industrial applicability
- IV ☐ Lack of unity of invention
- V ☒ Reasoned statement under Article 35(2) with regard to novelty, inventive step or industrial applicability; citations and explanations supporting such statement
- VI ☐ Certain documents cited
- VII ☐ Certain defects in the international application
- VIII ☒ Certain observations on the international application

|   |   |
|---|---|
| Date of submission of the demand<br><br><b>26/07/2000</b>   | Date of completion of this report<br><br><b>15.06.2001</b>                              |
| Name and mailing address of the international preliminary examining authority:<br><br> <b>European Patent Office - P.B. 5818 Patentlaan 2<br/>NL-2280 HV Rijswijk - Pays Bas<br/>Tel. +31 70 340 - 2040 Tx: 31 651 epo nl<br/>Fax: +31 70 340 - 3016</b> | Authorized officer<br><br><b>Macchia, G</b><br><br>Telephone No. <b>+31 70 340 4078</b> |



**INTERNATIONAL PRELIMINARY  
EXAMINATION REPORT**

International application No. PCT/EP00/00346

**I. Basis of the report**

1. With regard to the **elements** of the international application (*Replacement sheets which have been furnished to the receiving Office in response to an invitation under Article 14 are referred to in this report as "originally filed" and are not annexed to this report since they do not contain amendments (Rules 70.16 and 70.17)*):  
**Description, pages:**

1-37 as originally filed

**Claims, No.:**

1-5 with telefax of 27/04/2001

**Drawings, sheets:**

1/1 as originally filed

**Sequence listing part of the description, pages:**

1-5, filed with the letter of 28/03/2000

2. With regard to the **language**, all the elements marked above were available or furnished to this Authority in the language in which the international application was filed, unless otherwise indicated under this item.

These elements were available or furnished to this Authority in the following language: , which is:

- ☐ the language of a translation furnished for the purposes of the international search (under Rule 23.1(b)).  
☐ the language of publication of the international application (under Rule 48.3(b)).  
☐ the language of a translation furnished for the purposes of international preliminary examination (under Rule 55.2 and/or 55.3).

3. With regard to any **nucleotide and/or amino acid sequence** disclosed in the international application, the international preliminary examination was carried out on the basis of the sequence listing:

- ☐ contained in the international application in written form.  
☐ filed together with the international application in computer readable form.  
☒ furnished subsequently to this Authority in written form.  
☒ furnished subsequently to this Authority in computer readable form.  
☒ The statement that the subsequently furnished written sequence listing does not go beyond the disclosure in the international application as filed has been furnished.  
☒ The statement that the information recorded in computer readable form is identical to the written sequence listing has been furnished.

4. The amendments have resulted in the cancellation of:

# INTERNATIONAL PRELIMINARY EXAMINATION REPORT

International application No. PCT/EP00/00346

- ☐ the description, pages:
- ☐ the claims, Nos.:
- ☐ the drawings, sheets:

5. ☒ This report has been established as if (some of) the amendments had not been made, since they have been considered to go beyond the disclosure as filed (Rule 70.2(c)):

*(Any replacement sheet containing such amendments must be referred to under item 1 and annexed to this report.)*

**see separate sheet**

6. Additional observations, if necessary:

### III. Non-establishment of opinion with regard to novelty, inventive step and industrial applicability

1. The questions whether the claimed invention appears to be novel, to involve an inventive step (to be non-obvious), or to be industrially applicable have not been examined in respect of:

- ☐ the entire international application.
- ☒ claims Nos. 1-5, insofar as they relate to SEQ IDs NO:2-6.

because:

- ☐ the said international application, or the said claims Nos. relate to the following subject matter which does not require an international preliminary examination (*specify*):
- ☐ the description, claims or drawings (*indicate particular elements below*) or said claims Nos. are so unclear that no meaningful opinion could be formed (*specify*):
- ☐ the claims, or said claims Nos. are so inadequately supported by the description that no meaningful opinion could be formed.
- ☒ no international search report has been established for the said claims Nos. 1-5, insofar as they relate to SEQ IDs NO:2-6.

2. A meaningful international preliminary examination cannot be carried out due to the failure of the nucleotide and/or amino acid sequence listing to comply with the standard provided for in Annex C of the Administrative Instructions:

- ☐ the written form has not been furnished or does not comply with the standard.
- ☐ the computer readable form has not been furnished or does not comply with the standard.

### V. Reasoned statement under Article 35(2) with regard to novelty, inventive step or industrial applicability; citations and explanations supporting such statement

# INTERNATIONAL PRELIMINARY EXAMINATION REPORT

International application No. PCT/EP00/00346

---

## 1. Statement

|                               |      |                         |
|-------------------------------|------|-------------------------|
| Novelty (N)                   | Yes: | Claims 1, 2, 3(a), 4, 5 |
|                               | No:  | Claims 1, 3(b,c)        |
| Inventive step (IS)           | Yes: | Claims                  |
|                               | No:  | Claims 1-5              |
| Industrial applicability (IA) | Yes: | Claims 1-5              |
|                               | No:  | Claims                  |

## 2. Citations and explanations **see separate sheet**

## VIII. Certain observations on the international application

The following observations on the clarity of the claims, description, and drawings or on the question whether the claims are fully supported by the description, are made:  
**see separate sheet**

**INTERNATIONAL PRELIMINARY  
EXAMINATION REPORT - SEPARATE SHEET**

---

International application No. PCT/EP00/00346

**Re Item I**

**Basis of the report**

The amendments filed with the letter dated 27 April 2001 introduce subject-matter which extends beyond the content of the application as filed, contrary to Article 34(2)(b) PCT.

Amended claim 3 makes reference to a fragment of the polynucleotide of the invention, " said fragment (if necessary when coupled to a carrier) being capable of raising an immune response which recognises the polypeptide encoded by SEQ ID NO:1 ".

No support for this amendment can be found in present application as filed, in particular, on page 7, lines 5-7 of present application it is referred to a **polypeptide** fragment having a certain immunogenic activity, and **not** to a **polynucleotide** fragment. Consequently, claim 3 as originally filed has instead been taken into consideration.

**Re Item III**

**Non-establishment of opinion with regard to novelty, inventive step and industrial applicability**

No required additional search fees were timely paid by the applicant.

Therefore, no International Search Report has been established for the subject-matter related to the polynucleotides as in SEQ IDs NO:2-6.

Consequently, no opinion is established for the subject-matter related to the polynucleotides as in SEQ IDs NO:2-6.

**Re Item V**

**Reasoned statement under Article 35(2) with regard to novelty, inventive step or industrial applicability; citations and explanations supporting such statement**

1). Reference is made to the following documents:

D1: Nucleotide Sequence Database EMBL ID HS1220328, Accession number AA422178, 19 May 1997;

**INTERNATIONAL PRELIMINARY  
EXAMINATION REPORT - SEPARATE SHEET**

---

International application No. PCT/EP00/00346

D2: Nucleotide Sequence Database EMBL ID AA587764, Accession number AA587764, 18 September 1997;

D3: Nucleotide Sequence Database EMBL ID AI339648, Accession number AI339648, 31 December 1998.

- 2). Document D1 discloses a polynucleotide that is 99.8% identical, in reverse orientation, to the polynucleotide expressed as SEQ ID NO:1 of present application in the range from nucleotide 281 to nucleotide 784 of SEQ ID NO:1. Document D2 discloses a polynucleotide that is 99.3% identical to the polynucleotide expressed as SEQ ID NO:1 of present application in the range from nucleotide 1 to nucleotide 441 of SEQ ID NO:1. Document D3 discloses a polynucleotide that is 100% identical to the polynucleotide expressed as SEQ ID NO:1 of present application, in the range from nucleotide 6 to nucleotide 411 of SEQ ID NO:1. Therefore the polynucleotides disclosed in D1-D3 can be considered as fragments of a polynucleotide as claimed in claims 1 and 3(c). Moreover, said polynucleotides hybridize with the polynucleotide having the sequence of SEQ ID NO:1, therefore, they are embraced by the scope of claim 3(b). As a consequence, documents D2 and D3 are novelty destroying for claim 1 (as far as claim 1 refers to fragments of the polynucleotide as in SEQ ID NO:1) and documents D1, D2 and D3 are novelty destroying for claims 3(b) and 3(c) because there are no features in claims 1 (as far as claim 1 refers to fragments of the polynucleotide as in SEQ ID NO:1) and 3(b,c) that would allow the polynucleotides of said claims to be distinguished from the ones of documents D1-D3 (Article 33(2) PCT).
- 3). Claims 1 (as far as claim 1 refers to a polynucleotide comprising a nucleotide sequence which has at least 70% identity to that of SEQ ID NO:1, over the entire length of SEQ ID NO:1), 2, 3(a), 4 and 5 meet the requirements of Article 33(2) PCT because a polynucleotide comprising a nucleotide sequence which has at least 70% identity to that of SEQ ID NO:1, over the entire length of SEQ ID NO:1, and related diagnostic processes were not described in the available prior art (documents D1-D3).

- 4). The following remarks should be noted with reference to the inventivity of the polynucleotide as in SEQ ID NO:1 (Article 33(3) PCT).

The present application does not give any indication about the activity or functional role of the polynucleotide of SEQ ID NO:1 and of a possibly encoded polypeptide.

In fact, present application does not give any clear indication about the over-expression of the polynucleotide as in SEQ ID NO:1 in colon cancer tissues. Furthermore, the way the data are expressed in tables 1 and 3 of present application, does not allow the skilled person to identify the polynucleotide of SEQ ID NO:1 as a specific marker for colon cancer.

As a matter of fact, on page 30 of present application it is recited that " the difference between Ct values (CtN-CtT) is a direct measure of the difference in transcript levels between the tumour and normal tissues ". In table 1, 15 patients were considered relevant, over a **not specified** number of tested patients. However, it is not specified if the values under " colon range Ct " and " colon mean Ct " refer to a normal tissue, a tumour tissue or a ratio between the twos, and in this case in which order.

Moreover, in the last line of page 31 and at the beginning of page 32, it is stated that " 2 types of sequences are described, classified into categories A and B. A sequence of category A has a colon-specific expression. A sequence of category B is over-expressed in colon cancer ". The sequence having internal denomination " CASB611 " is classified in class A, therefore it should be considered a colon specific marker, not a colon cancer-specific marker.

Consequently since present application does not provide any support to the assumption that the polynucleotide as in SEQ ID NO:1 is over-expressed in colon cancer, its diagnostic role is purely speculative.

In addition to this, in present application it is referred in very general terms to a polypeptide encoded by the polynucleotide of SEQ ID NO:1, immunogenic fragments thereof, vaccines comprising the polynucleotide of SEQ ID NO:1 or its putative encoded polypeptide, a diagnostic kit comprising said polynucleotide or polypeptide, specific antibodies and ligands of said polypeptide, applications of the polynucleotide in chromosomal localization or applications of the putative encoded polypeptide in screening for agonists and antagonists of said polypeptide.

Again, no further specification is given about the above mentioned features and,

concerning the polypeptide encoded by the polynucleotide of present invention, the description does not provide any demonstration that it is actually produced, therefore its possible role as a candidate vaccine antigen is totally speculative. In the absence of a confirmed activity/function for the polynucleotide of SEQ ID NO:1 (or likewise for the polypeptide possibly encoded by said polynucleotide), the problem to be solved is reduced to the problem of provision of new polynucleotides.

This problem is not technically meaningful.

The solution is a non-inventive selection from a host of polynucleotides.

Consequently, the subject-matter of claims 1-5 does not involve an inventive step (Article 33(3) PCT).

- 5). The following remarks should be noted with reference to the industrial applicability of the polynucleotide as in SEQ ID NO:1 (Article 33(4) PCT).

In view of the lack of determined biological activity/function of the polynucleotide as in SEQ ID NO:1 and in view of its speculative uses, the industrial applicability does not appear to be plausible.

However, with regard to the different laws in the PCT contracting states, no objection is raised for the subject-matter of claims 1-5 on this issue.

#### **Re Item VIII**

##### **Certain observations on the international application**

- 1). The following remarks should be noted with respect to Art. 6, Rule 6 PCT.

- 2.1). In claim 1, it is not clear how a fragment of a polynucleotide can have the same immunogenic properties of the entire polynucleotide.

In other words, it is not clear how a fragment can have the same immunogenic properties of the parts of the polynucleotide not comprised in the fragment itself.

- 2.2). Claim 1 does not meet the requirements of Article 6 PCT in that the matter for which protection is sought is not clearly defined. The following functional statement " fragments which have the same immunogenic properties as the polynucleotide of SEQ ID NO:1 " does not enable the skilled person to determine



which technical features are necessary for the fragments concerned, in order to perform the stated function of being immunogenic.

- 3). In claim 3(b), the hybridization and washing conditions, corresponding to "stringent hybridization conditions " are not specified. This does not allow the skilled person to unambiguously identify the claimed matter.  
According to the PCT Preliminary Examination Guidelines, the meaning of the terms of a claim should, as far as possible, be clear for the person skilled in the art **from the wording of the claim alone** (PCT Guidelines C-III, 4.2).  
In this case, the skilled person is well aware that the term " stringent hybridization conditions " is not sufficient to define unambiguously the subject-matter concerned, because the term " stringent " is a term which does not give any indication about how to achieve these conditions. Even the definitions " high / moderately / low stringent hybridization conditions ", if ever used, would be ambiguously interpreted and would not meet the requirements of the PCT Guidelines C-IV, III-4.5.  
In addition to this, a functional limitation of the hybridizing polynucleotides is also missing.
- 4). In claim 3(c), the term " fragment " does not allow the skilled person to unambiguously identify the claimed matter because it is not specified neither the minimal length of said fragment, in terms of number of nucleotides, nor any functional limitation of said fragment.
- 5). In claims 4 and 5 it is not clear what the term " activity of a polynucleotide " refers to.

**Claims**

1. An isolated polynucleotide which comprises a nucleotide sequence which has at least 70% identity to that of SEQ ID Nos:1-6 over the entire length of SEQ ID Nos:1-6; or a nucleotide sequence complementary to said isolated polynucleotide; and fragments of the polynucleotides which when administered to a subject have the same immunogenic properties as the polynucleotides of SEQ ID Nos:1-6.
2. The isolated polynucleotide as defined in claim 1 in which the identity is at least 95% to SEQ ID Nos:1-6 over the entire length of SEQ ID Nos:1-6.
3. An isolated polynucleotide selected from:
  - (a) the polynucleotide of any one of SEQ ID Nos:1-6; and
  - (b) a polynucleotide obtainable by screening an appropriate library under stringent hybridization conditions with a labelled probe having the sequence of any one of SEQ ID Nos:1-6 or a fragment thereof; and
  - (c) a fragment of (a) or (b) said fragment (if necessary when coupled to a carrier) being capable of raising an immune response which recognises the polypeptide encoded by SEQ ID Nos:1-6;or a nucleotide sequence complementary to said isolated polynucleotide.
4. A process for diagnosing a disease or a susceptibility to a disease in a subject related to presence or expression or activity of a polynucleotide of any one of claims 1 to 3 in a subject comprising analyzing for the presence or amount of said polynucleotide in a sample derived from said subject.
5. A process for diagnosing the presence of colon cancer or a susceptibility to colon cancer in a subject related to presence or expression or activity of a polynucleotide of any one of claims 1 to 3 in a subject comprising analyzing for the presence or amount of said polynucleotide in a sample derived from said subject.

**PCT**WORLD INTELLECTUAL PROPERTY ORGANIZATION  
International Bureau

## INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

|  |                            |   |                            |    |           |                            |    |           |                            |    |           |                            |    |           |                            |    |           |                         |    |  |
|--|----------------------------|---|----------------------------|----|-----------|----------------------------|----|-----------|----------------------------|----|-----------|----------------------------|----|-----------|----------------------------|----|-----------|-------------------------|----|--|
| <b>(51) International Patent Classification 7 :</b><br><b>C12N 15/11, 15/12, C12Q 1/68</b>   | <b>A3</b>                  | <b>(11) International Publication Number:</b> <b>WO 00/43509</b><br><b>(43) International Publication Date:</b> 27 July 2000 (27.07.00) |                            |    |           |                            |    |           |                            |    |           |                            |    |           |                            |    |           |                         |    |  |
| <b>(21) International Application Number:</b> PCT/EP00/00346<br><b>(22) International Filing Date:</b> 17 January 2000 (17.01.00)<br><br><b>(30) Priority Data:</b> <table border="0"><tr><td>9901078.7</td><td>19 January 1999 (19.01.99)</td><td>GB</td></tr><tr><td>9902090.1</td><td>29 January 1999 (29.01.99)</td><td>GB</td></tr><tr><td>9902169.3</td><td>1 February 1999 (01.02.99)</td><td>GB</td></tr><tr><td>9902168.5</td><td>1 February 1999 (01.02.99)</td><td>GB</td></tr><tr><td>9902163.6</td><td>1 February 1999 (01.02.99)</td><td>GB</td></tr><tr><td>9907901.4</td><td>7 April 1999 (07.04.99)</td><td>GB</td></tr></table><br><b>(71) Applicant (for all designated States except US):</b> SMITHKLINE BEECHAM BIOLOGICALS S.A. [BE/BE]; Rue de l'Institut 89, B-1330 Rixensart (BE).<br><br><b>(72) Inventor; and</b><br><b>(75) Inventor/Applicant (for US only):</b> VINALS-BASSOLS, Carlota [BE/BE]; SmithKline Beecham Biologicals S.A., Rue de l'Institut 89, B-1330 Rixensart (BE).<br><br><b>(74) Agent:</b> TYRRELL, Arthur, William, Russell; SmithKline Beecham, Two New Horizons Court, Brentford, Middlesex TW8 9EP (GB). |                            | 9901078.7   | 19 January 1999 (19.01.99) | GB | 9902090.1 | 29 January 1999 (29.01.99) | GB | 9902169.3 | 1 February 1999 (01.02.99) | GB | 9902168.5 | 1 February 1999 (01.02.99) | GB | 9902163.6 | 1 February 1999 (01.02.99) | GB | 9907901.4 | 7 April 1999 (07.04.99) | GB | <b>(81) Designated States:</b> AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CR, CU, CZ, DE, DK, DM, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).<br><br><b>Published</b><br><i>With international search report.</i><br><br><b>(88) Date of publication of the international search report:</b><br>2 November 2000 (02.11.00) |
| 9901078.7  | 19 January 1999 (19.01.99) | GB  |                            |    |           |                            |    |           |                            |    |           |                            |    |           |                            |    |           |                         |    |  |
| 9902090.1  | 29 January 1999 (29.01.99) | GB  |                            |    |           |                            |    |           |                            |    |           |                            |    |           |                            |    |           |                         |    |  |
| 9902169.3  | 1 February 1999 (01.02.99) | GB  |                            |    |           |                            |    |           |                            |    |           |                            |    |           |                            |    |           |                         |    |  |
| 9902168.5  | 1 February 1999 (01.02.99) | GB  |                            |    |           |                            |    |           |                            |    |           |                            |    |           |                            |    |           |                         |    |  |
| 9902163.6  | 1 February 1999 (01.02.99) | GB  |                            |    |           |                            |    |           |                            |    |           |                            |    |           |                            |    |           |                         |    |  |
| 9907901.4  | 7 April 1999 (07.04.99)    | GB  |                            |    |           |                            |    |           |                            |    |           |                            |    |           |                            |    |           |                         |    |  |
| <b>(54) Title:</b> POLYNUCLEOTIDES AND THEIR APPLICATIONS IN DIAGNOSTICS<br><br><b>(57) Abstract</b><br><br>CASB611, CASB500, CASB501, CASB502, CASB505, or CASB507 polypeptides and polynucleotides and methods for producing such polypeptides by recombinant techniques are disclosed. Also disclosed are methods for utilizing CASB611, CASB500, CASB501, CASB502, CASB505, or CASB507 polypeptides and polynucleotides in diagnostics, and vaccines for prophylactic and therapeutic treatment of cancers, particularly colon cancers, autoimmune diseases, and related conditions.   |                            |   |                            |    |           |                            |    |           |                            |    |           |                            |    |           |                            |    |           |                         |    |  |

**FOR THE PURPOSES OF INFORMATION ONLY**

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

|    |                          |    |  |    |  |    |                          |
|----|--------------------------|----|--|----|--|----|--------------------------|
| AL | Albania                  | ES | Spain                                    | LS | Lesotho                                      | SI | Slovenia                 |
| AM | Armenia                  | FI | Finland                                  | LT | Lithuania                                    | SK | Slovakia                 |
| AT | Austria                  | FR | France                                   | LU | Luxembourg                                   | SN | Senegal                  |
| AU | Australia                | GA | Gabon                                    | LV | Latvia                                       | SZ | Swaziland                |
| AZ | Azerbaijan               | GB | United Kingdom                           | MC | Monaco                                       | TD | Chad                     |
| BA | Bosnia and Herzegovina   | GE | Georgia                                  | MD | Republic of Moldova                          | TG | Togo                     |
| BB | Barbados                 | GH | Ghana                                    | MG | Madagascar                                   | TJ | Tajikistan               |
| BE | Belgium                  | GN | Guinea                                   | MK | The former Yugoslav<br>Republic of Macedonia | TM | Turkmenistan             |
| BF | Burkina Faso             | GR | Greece                                   |    |  | TR | Turkey                   |
| BG | Bulgaria                 | HU | Hungary                                  | ML | Mali   | TT | Trinidad and Tobago      |
| BJ | Benin                    | IE | Ireland                                  | MN | Mongolia                                     | UA | Ukraine                  |
| BR | Brazil                   | IL | Israel                                   | MR | Mauritania                                   | UG | Uganda                   |
| BY | Belarus                  | IS | Iceland                                  | MW | Malawi                                       | US | United States of America |
| CA | Canada                   | IT | Italy                                    | MX | Mexico                                       | UZ | Uzbekistan               |
| CF | Central African Republic | JP | Japan                                    | NE | Niger  | VN | Viet Nam                 |
| CG | Congo                    | KE | Kenya                                    | NL | Netherlands                                  | YU | Yugoslavia               |
| CH | Switzerland              | KG | Kyrgyzstan                               | NO | Norway                                       | ZW | Zimbabwe                 |
| CI | Côte d'Ivoire            | KP | Democratic People's<br>Republic of Korea | NZ | New Zealand                                  |    |                          |
| CM | Cameroon                 | KR | Republic of Korea                        | PL | Poland                                       |    |                          |
| CN | China                    | KZ | Kazakhstan                               | PT | Portugal                                     |    |                          |
| CU | Cuba                     | LC | Saint Lucia                              | RO | Romania                                      |    |                          |
| CZ | Czech Republic           | LI | Liechtenstein                            | RU | Russian Federation                           |    |                          |
| DE | Germany                  | LK | Sri Lanka                                | SD | Sudan  |    |                          |
| DK | Denmark                  | LR | Liberia                                  | SE | Sweden                                       |    |                          |
| EE | Estonia                  |    |  | SG | Singapore                                    |    |                          |

# INTERNATIONAL SEARCH REPORT

International Application No  
PCT/EP 00/00346

|   |  |   |
|---|--|---|
| <b>A. CLASSIFICATION OF SUBJECT MATTER</b><br>IPC 7 C12N15/11 C12N15/12 C12Q1/68  |  |   |
| According to International Patent Classification (IPC) or to both national classification and IPC   |  |   |
| <b>B. FIELDS SEARCHED</b><br>Minimum documentation searched (classification system followed by classification symbols)<br>IPC 7 C12N C12Q   |  |   |
| Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched   |  |   |
| Electronic data base consulted during the international search (name of data base and, where practical, search terms used)  |  |   |
| <b>C. DOCUMENTS CONSIDERED TO BE RELEVANT</b>   |  |   |
| Category *  | Citation of document, with indication, where appropriate, of the relevant passages   | Relevant to claim No.   |
| X   | Nucleotide Sequence Database EMBL<br>ID HS1220328, Accession number AA422178<br>19 may 1997<br>99.8% identity with Seq.ID:1 nt.281-784<br>reverse orientation<br>XP002136186<br>the whole document   | 1,3   |
| X   | ---<br>Nucleotide Sequence Database EMBL<br>ID AA587764, Accession number AA587764<br>18 september 1997<br>99.3% identity with Seq.ID:1 nt.1-441<br>XP002136187<br>the whole document<br>---<br><div style="text-align: center;">-/-</div> | 1,3   |
| <div style="display: flex; justify-content: space-between;"> <span><input checked="" type="checkbox"/> Further documents are listed in the continuation of box C.</span> <span><input checked="" type="checkbox"/> Patent family members are listed in annex.</span> </div>   |  |   |
| <b>* Special categories of cited documents :</b>  |  |   |
| <div style="display: flex; justify-content: space-between;"> <div style="width: 45%;"> <p>"A" document defining the general state of the art which is not considered to be of particular relevance</p> <p>"E" earlier document but published on or after the international filing date</p> <p>"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)</p> <p>"O" document referring to an oral disclosure, use, exhibition or other means</p> <p>"P" document published prior to the international filing date but later than the priority date claimed</p> </div> <div style="width: 45%;"> <p>"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention</p> <p>"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone</p> <p>"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.</p> <p>"Z" document member of the same patent family</p> </div> </div> |  |   |
| Date of the actual completion of the international search<br><br><div style="text-align: center;">25 April 2000</div>   | Date of mailing of the international search report<br><br><div style="text-align: center;">13.07.00</div>  |   |
| Name and mailing address of the ISA<br>European Patent Office, P.B. 5818 Patentlaan 2<br>NL - 2280 HV Rijswijk<br>Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,<br>Fax: (+31-70) 340-3016  |  | Authorized officer<br><br><div style="text-align: center;">Macchia, G</div> |

# INTERNATIONAL SEARCH REPORT

International Application No  
PCT/cP 00/00346

| C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT |  |                       |
|--|--|-----------------------|
| Category *   | Citation of document, with indication, where appropriate, of the relevant passages   | Relevant to claim No. |
| X  | Nucleotide Sequence Database EMBL<br>ID A1339648. Accession number A1339648<br>31 december 1998<br>100% identity with Seq.ID:1 nt.6-411<br>XP002136188<br>the whole document | 1,3                   |
| A  | ---<br>WO 98 53319 A (THE JOHNS HOPKINS<br>UNIVERSITY (US); VOGELSTEIN BERT; KINZLER<br>KENNETH W.) 26 November 1998 (1998-11-26)<br>abstract<br>page 22 -page 80            | 4,5                   |
| A  | ---<br>WO 97 33903 A (SMITHKLINE BEECHAM CORP<br>(US); HUMAN GENOME SCIENCES INC (US); WEI<br>Y-F) 18 September 1997 (1997-09-18)<br>abstract<br>page 50; claim 19           | 4,5                   |
| A  | ---<br>WO 95 15381 A (JOHNS HOPKINS UNIV (US); DE<br>LA CHAPELLE A.; VOGELSTEIN B.; KINZLER<br>K.W) 8 June 1995 (1995-06-08)<br>abstract<br>page 48; claim 11<br>-----       | 4,5                   |

# INTERNATIONAL SEARCH REPORT

International application No.  
PCT/EP 00/00346

## Box I Observations where certain claims were found unsearchable (Continuation of Item 1 of first sheet)

This International Search Report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:  
because they relate to subject matter not required to be searched by this Authority, namely:
2. ☐ Claims Nos.:  
because they relate to parts of the International Application that do not comply with the prescribed requirements to such an extent that no meaningful International Search can be carried out, specifically:
3. ☐ Claims Nos.:  
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

## Box II Observations where unity of invention is lacking (Continuation of Item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

*see additional sheet*

1. ☐ As all required additional search fees were timely paid by the applicant, this International Search Report covers all searchable claims.
2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.
3. ☐ As only some of the required additional search fees were timely paid by the applicant, this International Search Report covers only those claims for which fees were paid, specifically claims Nos.:
4. ☒ No required additional search fees were timely paid by the applicant. Consequently, this International Search Report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

1-5 all partially

Remark on Protest

- ☐ The additional search fees were accompanied by the applicant's protest.
- ☐ No protest accompanied the payment of additional search fees.

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

1. Claims: 1-5 all partially

A polynucleotide comprising a nucleotide sequence which has at least 70% identity to that of Seq.ID:1, a nucleotide sequence complementary to said polynucleotide, or fragments of said polynucleotide. A polynucleotide as above, in which the identity to Seq.ID:1 is at least 95%. A polynucleotide selected from: the polynucleotide of Seq.ID:1, a polynucleotide obtainable by screening an appropriate library with a labeled probe having the sequence of Seq.ID:1 or a fragment thereof, fragments or complementary nucleotide sequences thereof. A process for diagnosing a disease or susceptibility of disease, namely colon cancer, related to the presence or expression or activity of a polynucleotide, comprising analyzing the presence or amount of said polynucleotide in a sample derived from a subject.

2. Claims: 1-5 all partially

As invention 1 but concerning Seq.ID:2

3. Claims: 1-5 all partially

As invention 1 but concerning Seq.ID:3

4. Claims: 1-5 all partially

As invention 1 but concerning Seq.ID:4

5. Claims: 1-5 all partially

As invention 1 but concerning Seq.ID:5

6. Claims: 1-5 all partially

As invention 1 but concerning Seq.ID:6



# INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No

PCT/EP 00/00346

| Patent document<br>cited in search report |   | Publication<br>date | Patent family<br>member(s) | Publication<br>date |
|---|---|---------------------|----------------------------|---------------------|
| WO 9853319                                | A | 26-11-1998          | AU 7499198 A               | 11-12-1998          |
| WO 9733903                                | A | 18-09-1997          | AU 5305596 A               | 01-10-1997          |
| WO 9515381                                | A | 08-06-1995          | EP 0730648 A               | 11-09-1996          |
|   |   |                     | JP 9506509 T               | 30-06-1997          |
|   |   |                     | US 5693470 A               | 02-12-1997          |
|   |   |                     | US 5837443 A               | 17-11-1998          |

PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION  
International Bureau



INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

|   |  |           |  |
|---|--|-----------|--|
| (51) International Patent Classification <sup>7</sup> :<br><b>C12N 15/11, 15/12, C12Q 1/68</b>  |  | <b>A2</b> | (11) International Publication Number: <b>WO 00/43509</b><br>(43) International Publication Date: <b>27 July 2000 (27.07.00)</b>   |
| (21) International Application Number: <b>PCT/EP00/00346</b><br>(22) International Filing Date: <b>17 January 2000 (17.01.00)</b><br><br>(30) Priority Data:<br>9901078.7      19 January 1999 (19.01.99)      GB<br>9902090.1      29 January 1999 (29.01.99)      GB<br>9902169.3      1 February 1999 (01.02.99)      GB<br>9902168.5      1 February 1999 (01.02.99)      GB<br>9902163.6      1 February 1999 (01.02.99)      GB<br>9907901.4      7 April 1999 (07.04.99)      GB           |  |           | (81) Designated States: AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CR, CU, CZ, DE, DK, DM, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG). |
| (71) Applicant (for all designated States except US): SMITHKLINE BEECHAM BIOLOGICALS S.A. [BE/BE]; Rue de l'Institut 89, B-1330 Rixensart (BE).<br><br>(72) Inventor; and<br>(75) Inventor/Applicant (for US only): VINALS-BASSOLS, Carlota [BE/BE]; SmithKline Beecham Biologicals S.A., Rue de l'Institut 89, B-1330 Rixensart (BE).<br><br>(74) Agent: TYRRELL, Arthur, William, Russell; SmithKline Beecham, Two New Horizons Court, Brentford, Middlesex TW8 9EP (GB).                       |  |           | Published<br>Without international search report and to be republished upon receipt of that report.  |
| (54) Title: <b>NOVEL COMPOUNDS</b>  |  |           |  |
| (57) Abstract<br><br>CASB611, CASB500, CASB501, CASB502, CASB505, or CASB507 polypeptides and polynucleotides and methods for producing such polypeptides by recombinant techniques are disclosed. Also disclosed are methods for utilizing CASB611, CASB500, CASB501, CASB502, CASB505, or CASB507 polypeptides and polynucleotides in diagnostics, and vaccines for prophylactic and therapeutic treatment of cancers, particularly colon cancers, autoimmune diseases, and related conditions. |  |           |  |

**FOR THE PURPOSES OF INFORMATION ONLY**

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

|    |                          |    |  |    |  |    |                          |
|----|--------------------------|----|--|----|--|----|--------------------------|
| AL | Albania                  | ES | Spain                                    | LS | Lesotho                                      | SI | Slovenia                 |
| AM | Armenia                  | FI | Finland                                  | LT | Lithuania                                    | SK | Slovakia                 |
| AT | Austria                  | FR | France                                   | LU | Luxembourg                                   | SN | Senegal                  |
| AU | Australia                | GA | Gabon                                    | LV | Latvia                                       | SZ | Swaziland                |
| AZ | Azerbaijan               | GB | United Kingdom                           | MC | Monaco                                       | TD | Chad                     |
| BA | Bosnia and Herzegovina   | GE | Georgia                                  | MD | Republic of Moldova                          | TG | Togo                     |
| BB | Barbados                 | GH | Ghana                                    | MG | Madagascar                                   | TJ | Tajikistan               |
| BE | Belgium                  | GN | Guinea                                   | MK | The former Yugoslav<br>Republic of Macedonia | TM | Turkmenistan             |
| BF | Burkina Faso             | GR | Greece                                   | ML | Mali   | TR | Turkey                   |
| BG | Bulgaria                 | HU | Hungary                                  | MN | Mongolia                                     | TT | Trinidad and Tobago      |
| BJ | Benin                    | IE | Ireland                                  | MR | Mauritania                                   | UA | Ukraine                  |
| BR | Brazil                   | IL | Israel                                   | MW | Malawi                                       | UG | Uganda                   |
| BY | Belarus                  | IS | Iceland                                  | MX | Mexico                                       | US | United States of America |
| CA | Canada                   | IT | Italy                                    | NE | Niger  | UZ | Uzbekistan               |
| CF | Central African Republic | JP | Japan                                    | NL | Netherlands                                  | VN | Viet Nam                 |
| CG | Congo                    | KE | Kenya                                    | NO | Norway                                       | YU | Yugoslavia               |
| CH | Switzerland              | KG | Kyrgyzstan                               | NZ | New Zealand                                  | ZW | Zimbabwe                 |
| CI | Côte d'Ivoire            | KP | Democratic People's<br>Republic of Korea | PL | Poland                                       |    |                          |
| CM | Cameroon                 | KR | Republic of Korea                        | PT | Portugal                                     |    |                          |
| CN | China                    | KZ | Kazakhstan                               | RO | Romania                                      |    |                          |
| CU | Cuba                     | LC | Saint Lucia                              | RU | Russian Federation                           |    |                          |
| CZ | Czech Republic           | LI | Liechtenstein                            | SD | Sudan  |    |                          |
| DE | Germany                  | LK | Sri Lanka                                | SE | Sweden                                       |    |                          |
| DK | Denmark                  | LR | Liberia                                  | SG | Singapore                                    |    |                          |
| EE | Estonia                  |    |  |    |  |    |                          |

WO 00/43509

PCT/EP00/00346

## Novel Compounds

### FIELD OF THE INVENTION

The present invention relates to polynucleotides, (herein referred to as "CASB611  
 5 polynucleotide(s)", "CASB500 polynucleotide(s)", "CASB501 polynucleotide(s)",  
 "CASB502 polynucleotide(s)", "CASB505 polynucleotide(s)", "CASB507  
 polynucleotide(s)"), polypeptides encoded thereby (referred to herein as "CASB611",  
 "CASB500", "CASB501", "CASB502", "CASB505", and "CASB507" respectively or  
 "CASB611 polypeptide(s)", "CASB500 polypeptide (s)", "CASB501 polypeptide (s)",  
 10 "CASB502 polypeptide (s)", "CASB505 polypeptide (s)", "CASB507 polypeptide  
 (s)" respectively), recombinant materials and methods for their production. In another  
 aspect, the invention relates to methods for using such polypeptides and polynucleotides,  
 including the diagnostics and treatment of cancer and autoimmune diseases and other related  
 conditions. In a further aspect, the invention relates to methods for identifying agonists  
 15 and antagonists/inhibitors using the materials provided by the invention, and treating  
 conditions associated with CASB611, CASB500, CASB501, CASB502, CASB505, or  
 CASB507 polypeptide imbalance with the identified compounds. In a still further aspect,  
 the invention relates to diagnostic assays for detecting diseases associated with inappropriate  
 CASB611, CASB500, CASB501, CASB502, CASB505, or CASB507 polypeptide  
 20 activity or levels.

### BACKGROUND OF THE INVENTION

Polynucleotides and polypeptides of the present invention are believed to be important  
 immunogens for specific prophylactic or therapeutic immunization against tumours, because  
 25 they are specifically expressed or highly over-expressed in tumours compared to normal  
 cells and can thus be targeted by antigen-specific immune mechanisms leading to the  
 destruction of the tumour cell. They can also be used to diagnose the occurrence of tumour  
 cells. Furthermore, their inappropriate expression in certain circumstances can cause an  
 induction of autoimmune, inappropriate immune responses, which could be corrected  
 30 through appropriate vaccination using the same polypeptides or polynucleotides. In this  
 respect the most important biological activities to our purpose are the antigenic and  
 immunogenic activities of the polypeptide of the present invention. A polypeptide of the  
 present invention may also exhibit at least one other biological activity of a CASB611,

CASB500, CASB501, CASB502, CASB505, or CASB507 polypeptide, which could qualify it as a target for therapeutic or prophylactic intervention different from that linked to the immune response.

5 Functional genomics relies heavily on high-throughput DNA sequencing technologies and the various tools of bioinformatics to identify gene sequences of potential interest from the many molecular biology databases now available. cDNA libraries enriched for genes of relevance to a particular tissue or physiological situation can be constructed using recently developed subtractive cloning strategies. Furthermore, cDNAs found in libraries of certain  
10 tissues and not others can be identified using appropriate electronic screening methods.

High throughput genome- or gene-based biology allows new approaches to the identification and cloning of target genes for useful immune responses for the prevention and vaccine therapy of diseases such as cancer and autoimmunity.

15

## DESCRIPTION OF THE INVENTION

### Polynucleotides

In a first aspect, the present invention relates to CASB611, CASB500, CASB501, CASB502, CASB505, or CASB507 polynucleotides. Such polynucleotides include  
20 isolated polynucleotides comprising a nucleotide sequence which has at least 70% identity, preferably at least 80% identity, more preferably at least 90% identity, yet more preferably at least 95% identity, to SEQ ID NOs:1 - 6 respectively over the entire length of SEQ ID NOs:1 - 6. In this regard, polynucleotides which have at least 97% identity are highly preferred, whilst those with at least 98-99% identity are more highly preferred, and  
25 those with at least 99% identity are most highly preferred. Such polynucleotides include a polynucleotide comprising the polynucleotide of SEQ ID NOs:1 - 6 as well as the polynucleotide of SEQ ID NOs:1 - 6. Said polynucleotide can be inserted in a suitable plasmid or recombinant microorganism vector and used for immunization ( see for example Wolff et. al., Science 247:1465-1468 (1990); Corr et. al., J. Exp. Med. 184:1555-1560  
30 (1996); Doe et. al., Proc. Natl. Acad. Sci. 93:8578-8583 (1996)).

The invention also provides polynucleotides which are complementary to the above described polynucleotides.

The invention also provides a polynucleotide comprising a nucleotide sequence that has at least 70% identity to a nucleotide sequence encoding a polypeptide of the invention over the entire coding region, encoded by a polynucleotide comprising the sequence contained in any one of SEQ ID Nos: 1 - 6; or a nucleotide sequence complimentary to said isolated  
5 polynucleotide.

The invention also provides a fragment of a CASB611, CASB500, CASB501, CASB502, CASB505, or CASB507 polynucleotide which when administered to a subject has the same immunogenic properties as the polynucleotide of SEQ ID NOs:1 - 6.

10

The invention also provides a polynucleotide encoding an immunological fragment of a CASB611, CASB500, CASB501, CASB502, CASB505, or CASB507 polypeptide as hereinbefore defined.

15 The nucleotide sequences of SEQ ID NO:1 (CASB611) and SEQ ID NO:2 show no homology with any known gene.. The nucleotide sequences of SEQ ID NO:3-6 show homology to intron portions of the chromosomal PBEF gene (Homo sapiens BAC clone RP11-22N19 from 7q22, complete sequence; ACCESSION : AC007032), SEQ ID NO: 3 and 5 (CASB501 and CASB505) being located between exon 8 and exon 7 of the PBEF  
20 gene (nucleotides 67 484 to 67 966 and 68 107 to 68 583), and SEQ ID NO:4 (CASB502) and 6 (CASB507) being located downstream of exon 8 (from 69 274 to 69 732 and from 70 163 to 71 223).

CASB500 is located on chromosome 6 (accession number HSJ651N20).

25

Preferred polypeptides and polynucleotides of the present invention are expected to have, *inter alia*, similar biological functions/properties to their homologous polypeptides and polynucleotides. Furthermore, preferred polynucleotides of the present invention have at least one activity of SEQ ID NOs:1 - 6, as appropriate.

30

The present invention also relates to partial polynucleotide and polypeptide sequences which were first identified prior to the determination of the corresponding sequences of SEQ ID NOs:1 - 6.

Accordingly, in a further aspect, the present invention provides for an isolated polynucleotide which:

- 5 (a) comprises a nucleotide sequence which has at least 70% identity, preferably at least 80% identity, more preferably at least 90% identity, yet more preferably at least 95% identity, even more preferably at least 97-99% identity to SEQ ID NO:7-12 over the entire length of SEQ ID NO:7-12;
- (b) has a nucleotide sequence which has at least 70% identity, preferably at least 80% identity, more preferably at least 90% identity, yet more preferably at least 95% identity, 10 even more preferably at least 97-99% identity, to SEQ ID NO:1-6 over the entire length of SEQ ID NO:7-12; or
- (c) the polynucleotides of SEQ ID NO:7-12; or

15 The nucleotide sequences of SEQ ID NO:7-12 and the peptide sequence encoded thereby are derived from EST (Expressed Sequence Tag) sequences. It is recognised by those skilled in the art that there will inevitably be some nucleotide sequence reading errors in EST sequences (see Adams, M.D. *et al*, Nature 377 (supp) 3, 1995). Accordingly, the nucleotide sequence of SEQ ID NO:7-12 and the peptide sequence encoded therefrom are therefore subject to the same inherent limitations in sequence accuracy.

20 Polynucleotides of the present invention may be obtained, using standard cloning and screening techniques, from a cDNA library derived from mRNA in cells of human colon cancer, (for example Sambrook et al., Molecular Cloning: A Laboratory Manual, 2<sup>nd</sup> Ed., Cold Spring harbor Laboratory Press, Cold Spring harbor, N.Y. (1989)). Polynucleotides of 25 the invention can also be obtained from natural sources such as genomic DNA libraries or can be synthesized using well known and commercially available techniques.

When polynucleotides of the present invention are used for the recombinant production of polypeptides of the present invention, the polynucleotide may include the coding sequence 30 for the mature polypeptide, by itself; or the coding sequence for the mature polypeptide in reading frame with other coding sequences, such as those encoding a leader or secretory sequence, a pre-, or pro- or prepro- protein sequence, or other fusion peptide portions. For example, a marker sequence which facilitates purification of the fused polypeptide can be

encoded. In certain preferred embodiments of this aspect of the invention, the marker sequence is a hexa-histidine peptide, as provided in the pQE vector (Qiagen, Inc.) and described in Gentz *et al.*, Proc Natl Acad Sci USA (1989) 86:821-824, or is an HA tag. The polynucleotide may also contain non-coding 5' and 3' sequences, such as transcribed, non-  
5 translated sequences, splicing and polyadenylation signals, ribosome binding sites and sequences that stabilize mRNA.

Polynucleotides which are identical or sufficiently identical to a nucleotide sequence contained in SEQ ID NOs:1 - 6, may be used as hybridization probes for cDNA and  
10 genomic DNA or as primers for a nucleic acid amplification (PCR) reaction, to isolate full-length cDNAs and genomic clones encoding polypeptides of the present invention and to isolate cDNA and genomic clones of other genes (including genes encoding paralogs from human sources and orthologs and paralogs from species other than human) that have a high sequence similarity to SEQ ID NO:1 - 6. Typically these nucleotide sequences are 70%  
15 identical, preferably 80% identical, more preferably 90% identical, most preferably 95% identical to that of the referent. The probes or primers will generally comprise at least 15 nucleotides, preferably, at least 30 nucleotides and may have at least 50 nucleotides. Particularly preferred probes will have between 30 and 50 nucleotides. Particularly preferred primers will have between 20 and 25 nucleotides. In particular, polypeptides or  
20 polynucleotides derived from sequences from homologous animal origin could be used as immunogens to obtain a cross-reactive immune response to the human gene.

A polynucleotide encoding a polypeptide of the present invention, including homologs from species other than human, may be obtained by a process which comprises the steps of  
25 screening an appropriate library under stringent hybridization conditions with a labeled probe having the sequence of SEQ ID NOs: 1 - 6 or a fragment thereof; and isolating full-length cDNA and genomic clones containing said polynucleotide sequence. Such hybridization techniques are well known to the skilled artisan. Preferred stringent hybridization conditions include overnight incubation at 42°C in a solution comprising: 50%  
30 formamide, 5xSSC (150mM NaCl, 15mM trisodium citrate), 50 mM sodium phosphate (pH7.6), 5x Denhardt's solution, 10 % dextran sulfate, and 20 microgram/ml denatured, sheared salmon sperm DNA; followed by washing the filters in 0.1x SSC at about 65°C.



Thus the present invention also includes polynucleotides obtainable by screening an appropriate library under stringent hybridization conditions with a labeled probe having the sequence of SEQ ID NOs:1 – 6 or a fragment thereof.

- 5 The skilled artisan will appreciate that, in many cases, an isolated cDNA sequence will be incomplete, in that the region coding for the polypeptide is short at the 5' end of the cDNA.

There are several methods available and well known to those skilled in the art to obtain  
10 full-length cDNAs, or extend short cDNAs, for example those based on the method of Rapid Amplification of cDNA ends (RACE) (see, for example, Frohman et al., PNAS USA 85, 8998-9002, 1988). Recent modifications of the technique, exemplified by the Marathon™ technology (Clontech Laboratories Inc.) for example, have significantly  
15 simplified the search for longer cDNAs. In the Marathon™ technology, cDNAs have been prepared from mRNA extracted from a chosen tissue and an 'adaptor' sequence ligated onto each end. Nucleic acid amplification (PCR) is then carried out to amplify the 'missing' 5' end of the cDNA using a combination of gene specific and adaptor specific oligonucleotide primers. The PCR reaction is then repeated using 'nested' primers, that is,  
20 primers designed to anneal within the amplified product (typically an adaptor specific primer that anneals further 3' in the adaptor sequence and a gene specific primer that anneals further 5' in the known gene sequence). The products of this reaction can then be analysed by DNA sequencing and a full-length cDNA constructed either by joining the product directly to the existing cDNA to give a complete sequence, or carrying out a separate full-length PCR using the new sequence information for the design of the 5'  
25 primer.

### **Polypeptides**

In a further aspect, the present invention relates to CASB611, CASB500, CASB501, CASB502, CASB505, or CASB507 polypeptides.

30

Further peptides of the present invention include isolated polypeptides encoded by a polynucleotide comprising the sequence contained in one of SEQ ID NOS:1 - 6.

The invention also provides an immunogenic fragment of a CASB611, CASB500, CASB501, CASB502, CASB505, or CASB507 polypeptide, that is a contiguous portion of the CASB611, CASB500, CASB501, CASB502, CASB505, or CASB507 polypeptide which has the same or similar immunogenic properties to the polypeptide encoded by a polynucleotide comprising the sequence contained in one of SEQ ID NOS:1 – 6. That is to say, the fragment (if necessary when coupled to a carrier) is capable of raising an immune response which recognises the CASB611, CASB500, CASB501, CASB502, CASB505, or CASB507 polypeptide. Such an immunogenic fragment may include, for example, the CASB611, CASB500, CASB501, CASB502, CASB505, or CASB507 polypeptide lacking an N-terminal leader sequence, a transmembrane domain or a C-terminal anchor domain.

The polypeptides or immunogenic fragment of the invention may be in the form of the “mature” protein or may be a part of a larger protein such as a precursor or a fusion protein. It is often advantageous to include an additional amino acid sequence which contains secretory or leader sequences, pro-sequences, sequences which aid in purification such as multiple histidine residues, or an additional sequence for stability during recombinant production. Furthermore, addition of exogenous polypeptide or lipid tail or polynucleotide sequences to increase the immunogenic potential of the final molecule is also considered.

In one aspect, the invention relates to genetically engineered soluble fusion proteins comprising a polypeptide of the present invention, or a fragment thereof, and various portions of the constant regions of heavy or light chains of immunoglobulins of various subclasses (IgG, IgM, IgA, IgE). Preferred as an immunoglobulin is the constant part of the heavy chain of human IgG, particularly IgG1, where fusion takes place at the hinge region. In a particular embodiment, the Fc part can be removed simply by incorporation of a cleavage sequence which can be cleaved with blood clotting factor Xa. Furthermore, this invention relates to processes for the preparation of these fusion proteins by genetic engineering, and to the use thereof for drug screening, diagnosis and therapy. A further aspect of the invention also relates to polynucleotides encoding such fusion proteins. Examples of fusion protein technology can be found in International Patent Application Nos. WO94/29458 and WO94/22914.

The proteins may be chemically conjugated, or expressed as recombinant fusion proteins allowing increased levels to be produced in an expression system as compared to non-fused protein. The fusion partner may assist in providing T helper epitopes (immunological fusion partner), preferably T helper epitopes recognised by humans, or assist in expressing the protein (expression enhancer) at higher yields than the native recombinant protein. Preferably the fusion partner will be both an immunological fusion partner and expression enhancing partner.

Fusion partners include protein D from *Haemophilus influenza* B and the non-structural protein from influenzae virus, NS1 (hemagglutinin). Another immunological fusion partner is the protein known as LYTA. Preferably the C terminal portion of the molecule is used. Lyta is derived from *Streptococcus pneumoniae* which synthesize an N-acetyl-L-alanine amidase, amidase LYTA, (coded by the *lytA* gene {Gene, 43 (1986) page 265-272} an autolysin that specifically degrades certain bonds in the peptidoglycan backbone. The C-terminal domain of the LYTA protein is responsible for the affinity to the choline or to some choline analogues such as DEAE. This property has been exploited for the development of E.coli C-LYTA expressing plasmids useful for expression of fusion proteins. Purification of hybrid proteins containing the C-LYTA fragment at its amino terminus has been described {Biotechnology: 10, (1992) page 795-798}. It is possible to use the repeat portion of the Lyta molecule found in the C terminal end starting at residue 178, for example residues 188 - 305.

The present invention also includes variants of the aforementioned polypeptides, that is polypeptides that vary from the referents by conservative amino acid substitutions, whereby a residue is substituted by another with like characteristics. Typical such substitutions are among Ala, Val, Leu and Ile; among Ser and Thr; among the acidic residues Asp and Glu; among Asn and Gln; and among the basic residues Lys and Arg; or aromatic residues Phe and Tyr. Particularly preferred are variants in which several, 5-10, 1-5, 1-3, 1-2 or 1 amino acids are substituted, deleted, or added in any combination.

Polypeptides of the present invention can be prepared in any suitable manner. Such polypeptides include isolated naturally occurring polypeptides, recombinantly produced

polypeptides, synthetically produced polypeptides, or polypeptides produced by a combination of these methods. Means for preparing such polypeptides are well understood in the art.

5

#### **Vectors, Host cells, Expression Systems**

Recombinant polypeptides of the present invention may be prepared by processes well known in the art from genetically engineered host cells comprising expression systems. Accordingly, in a further aspect, the present invention relates to an expression system which  
10 comprises a polynucleotide of the present invention, to host cells which are genetically engineered with such expression systems and to the production of polypeptides of the invention by recombinant techniques. Cell-free translation systems can also be employed to produce such proteins using RNAs derived from the DNA constructs of the present invention.

15

For recombinant production, host cells can be genetically engineered to incorporate expression systems or portions thereof for polynucleotides of the present invention. Introduction of polynucleotides into host cells can be effected by methods described in many standard laboratory manuals, such as Davis et al., Basic Methods in Molecular Biology  
20 (1986) and Sambrook *et al.*, Molecular Cloning: A Laboratory Manual, 2nd Ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y. (1989). Preferred such methods include, for instance, calcium phosphate transfection, DEAE-dextran mediated transfection, transfection, microinjection, cationic lipid-mediated transfection, electroporation, transduction, scrape loading, ballistic introduction or infection.

25

Preferably the proteins of the invention are coexpressed with thioredoxin in trans (TIT). Coexpression of thioredoxin in trans versus in cis is preferred to keep antigen free of thioredoxin without the need for protease. Thioredoxin coexpression eases the solubilisation of the proteins of the invention. Thioredoxin coexpression has also a  
30 significant impact on protein purification yield, on purified-protein solubility and quality.

Representative examples of appropriate hosts include bacterial cells, such as *Streptococci*, *Staphylococci*, *E. coli*, *Streptomyces* and *Bacillus subtilis* cells; fungal cells, such as yeast

cells and *Aspergillus* cells; insect cells such as *Drosophila* S2 and *Spodoptera* Sf9 cells; animal cells such as CHO, COS, HeLa, C127, 3T3, BHK, HEK 293 and Bowes melanoma cells; and plant cells.

- 5 A great variety of expression systems can be used, for instance, chromosomal, episomal and virus-derived systems, e.g., vectors derived from bacterial plasmids, from bacteriophage, from transposons, from yeast episomes, from insertion elements, from yeast chromosomal elements, from viruses such as baculoviruses, papova viruses, such as SV40, vaccinia viruses, adenoviruses, fowl pox viruses, pseudorabies viruses and
- 10 retroviruses, and vectors derived from combinations thereof, such as those derived from plasmid and bacteriophage genetic elements, such as cosmids and phagemids. The expression systems may contain control regions that regulate as well as engender expression. Generally, any system or vector which is able to maintain, propagate or express a polynucleotide to produce a polypeptide in a host may be used. The appropriate
- 15 nucleotide sequence may be inserted into an expression system by any of a variety of well-known and routine techniques, such as, for example, those set forth in Sambrook *et al.*, Molecular Cloning, A Laboratory Manual (supra). Appropriate secretion signals may be incorporated into the desired polypeptide to allow secretion of the translated protein into the lumen of the endoplasmic reticulum, the periplasmic space or the extracellular
- 20 environment. These signals may be endogenous to the polypeptide or they may be heterologous signals.

- The expression system may also be a recombinant live microorganism, such as a virus or bacterium. The gene of interest can be inserted into the genome of a live recombinant
- 25 virus or bacterium. Inoculation and *in vivo* infection with this live vector will lead to *in vivo* expression of the antigen and induction of immune responses. Viruses and bacteria used for this purpose are for instance: poxviruses (e.g; vaccinia, fowlpox, canarypox), alphaviruses (Sindbis virus, Semliki Forest Virus, Venezuelan Equine Encephalitis Virus), adenoviruses, adeno-associated virus, picornaviruses (poliovirus, rhinovirus),
- 30 herpesviruses (varicella zoster virus, etc), *Listeria*, *Salmonella*, *Shigella*, BCG. These viruses and bacteria can be virulent, or attenuated in various ways in order to obtain live vaccines. Such live vaccines also form part of the invention.

Polypeptides of the present invention can be recovered and purified from recombinant cell cultures by well-known methods including ammonium sulfate or ethanol precipitation, acid extraction, anion or cation exchange chromatography, phosphocellulose chromatography, hydrophobic interaction chromatography, affinity chromatography, hydroxylapatite chromatography and lectin chromatography. Most preferably, ion metal affinity chromatography (IMAC) is employed for purification. Well known techniques for refolding proteins may be employed to regenerate active conformation when the polypeptide is denatured during intracellular synthesis, isolation and or purification.

#### 10 **Vaccines**

Another aspect of the invention relates to a method for inducing , re-inforcing or modulating an immunological response in a mammal which comprises inoculating the mammal with a fragment or the entire polypeptide or polynucleotide of the invention, adequate to produce antibody and/or T cell immune response for prophylaxis or for therapeutic treatment of cancer and autoimmune disease and related conditions. Yet another aspect of the invention relates to a method of inducing, re-inforcing or modulating immunological response in a mammal which comprises, delivering a polypeptide of the present invention *via* a vector or cell directing expression of the polynucleotide and coding for the polypeptide *in vivo* in order to induce such an immunological response to produce immune responses for prophylaxis or treatment of said mammal from diseases.

A further aspect of the invention relates to an immunological/vaccine formulation (composition) which, when introduced into a mammalian host, induces, re-inforces or modulates an immunological response in that mammal to a polypeptide of the present invention wherein the composition comprises a polypeptide or polynucleotide of the invention or an immunological fragment thereof as herein before defined. The vaccine formulation may further comprise a suitable carrier. Since a polypeptide may be broken down in the stomach, it is preferably administered parenterally (for instance, subcutaneous, intramuscular, intravenous, or intradermal injection). Formulations suitable for parenteral administration include aqueous and non-aqueous sterile injection solutions which may contain anti-oxidants, buffers, bacteriostats and solutes which render the formulation isotonic with the blood of the recipient; and aqueous and non-aqueous

sterile suspensions which may include suspending agents or thickening agents. The formulations may be presented in unit-dose or multi-dose containers, for example, sealed ampoules and vials and may be stored in a freeze-dried condition requiring only the addition of the sterile liquid carrier immediately prior to use.

5

A further aspect of the invention relates to the *in vitro* induction of immune responses to a fragment or the entire polypeptide or polynucleotide of the present invention or a molecule comprising the polypeptide or polynucleotide of the present invention, using cells from the immune system of a mammal, and reinfusing these activated immune cells of the mammal for the treatment of disease. Activation of the cells from the immune system is achieved by *in vitro* incubation with the entire polypeptide or polynucleotide of the present invention or a molecule comprising the polypeptide or polynucleotide of the present invention in the presence or absence of various immunomodulator molecules.

A further aspect of the invention relates to the immunization of a mammal by administration of antigen presenting cells modified by *in vitro* loading with part or the entire polypeptide of the present invention or a molecule comprising the polypeptide of the present invention and administered *in vivo* in an immunogenic way. Alternatively, antigen presenting cells can be transfected *in vitro* with a vector containing a fragment or the entire polynucleotide of the present invention or a molecule comprising the polynucleotide of the present invention, such as to express the corresponding polypeptide, and administered *in vivo* in an immunogenic way.

The vaccine formulation of the invention may also include adjuvant systems for enhancing the immunogenicity of the formulation. Preferably the adjuvant system raises preferentially a TH1 type of response.

An immune response may be broadly distinguished into two extreme categories, being a humoral or cell mediated immune responses (traditionally characterised by antibody and cellular effector mechanisms of protection respectively). These categories of response have been termed TH1-type responses (cell-mediated response), and TH2-type immune responses (humoral response).

Extreme TH1-type immune responses may be characterised by the generation of antigen specific, haplotype restricted cytotoxic T lymphocytes, and natural killer cell responses. In mice TH1-type responses are often characterised by the generation of antibodies of the IgG2a subtype, whilst in the human these correspond to IgG1 type antibodies. TH2-type immune responses are characterised by the generation of a broad range of immunoglobulin isotypes including in mice IgG1, IgA, and IgM.

It can be considered that the driving force behind the development of these two types of immune responses are cytokines. High levels of TH1-type cytokines tend to favour the induction of cell mediated immune responses to the given antigen, whilst high levels of TH2-type cytokines tend to favour the induction of humoral immune responses to the antigen.

The distinction of TH1 and TH2-type immune responses is not absolute. In reality an individual will support an immune response which is described as being predominantly TH1 or predominantly TH2. However, it is often convenient to consider the families of cytokines in terms of that described in murine CD4 +ve T cell clones by Mosmann and Coffman (*Mosmann, T.R. and Coffman, R.L. (1989) TH1 and TH2 cells: different patterns of lymphokine secretion lead to different functional properties. Annual Review of Immunology, 7, p145-173*). Traditionally, TH1-type responses are associated with the production of the INF- $\gamma$  and IL-2 cytokines by T-lymphocytes. Other cytokines often directly associated with the induction of TH1-type immune responses are not produced by T-cells, such as IL-12. In contrast, TH2- type responses are associated with the secretion of IL-4, IL-5, IL-6 and IL-13.

It is known that certain vaccine adjuvants are particularly suited to the stimulation of either TH1 or TH2 - type cytokine responses. Traditionally the best indicators of the TH1:TH2 balance of the immune response after a vaccination or infection includes direct measurement of the production of TH1 or TH2 cytokines by T lymphocytes *in vitro* after restimulation with antigen, and/or the measurement of the IgG1:IgG2a ratio of antigen specific antibody responses.



Thus, a TH1-type adjuvant is one which preferentially stimulates isolated T-cell populations to produce high levels of TH1-type cytokines when re-stimulated with antigen *in vitro*, and promotes development of both CD8+ cytotoxic T lymphocytes and antigen specific immunoglobulin responses associated with TH1-type isotype.

5

Adjuvants which are capable of preferential stimulation of the TH1 cell response are described in International Patent Application No. WO 94/00153 and WO 95/17209.

3 De-O-acylated monophosphoryl lipid A (3D-MPL) is one such adjuvant. This is known from GB 2220211 (Ribi). Chemically it is a mixture of 3 De-O-acylated monophosphoryl lipid A with 4, 5 or 6 acylated chains and is manufactured by Ribi Immunochem, Montana. A preferred form of 3 De-O-acylated monophosphoryl lipid A is disclosed in European Patent 0 689 454 B1 (SmithKline Beecham Biologicals SA).

15 Preferably, the particles of 3D-MPL are small enough to be sterile filtered through a 0.22micron membrane (European Patent number 0 689 454).

3D-MPL will be present in the range of 10µg - 100µg preferably 25-50µg per dose wherein the antigen will typically be present in a range 2-50µg per dose.

20 Another preferred adjuvant comprises QS21, an Hplc purified non-toxic fraction derived from the bark of Quillaja Saponaria Molina. Optionally this may be admixed with 3 De-O-acylated monophosphoryl lipid A (3D-MPL), optionally together with an carrier.

The method of production of QS21 is disclosed in US patent No. 5,057,540.

25

Non-reactogenic adjuvant formulations containing QS21 have been described previously (WO 96/33739). Such formulations comprising QS21 and cholesterol have been shown to be successful TH1 stimulating adjuvants when formulated together with an antigen.

30 Further adjuvants which are preferential stimulators of TH1 cell response include immunomodulatory oligonucleotides, for example unmethylated CpG sequences as disclosed in WO 96/02555.

Combinations of different TH1 stimulating adjuvants, such as those mentioned hereinabove, are also contemplated as providing an adjuvant which is a preferential stimulator of TH1 cell response. For example, QS21 can be formulated together with 3D-MPL. The ratio of QS21 : 3D-MPL will typically be in the order of 1 : 10 to 10 : 1; preferably 1:5 to 5 : 1 and often substantially 1 : 1. The preferred range for optimal synergy is 2.5 : 1 to 1 : 1 3D-MPL: QS21.

Preferably a carrier is also present in the vaccine composition according to the invention. The carrier may be an oil in water emulsion, or an aluminium salt, such as aluminium phosphate or aluminium hydroxide.

A preferred oil-in-water emulsion comprises a metabolisable oil, such as squalene, alpha tocopherol and Tween 80. In a particularly preferred aspect the antigens in the vaccine composition according to the invention are combined with QS21 and 3D-MPL in such an emulsion. Additionally the oil in water emulsion may contain span 85 and/or lecithin and/or tricaprylin.

Typically for human administration QS21 and 3D-MPL will be present in a vaccine in the range of 1µg - 200µg, such as 10-100µg, preferably 10µg - 50µg per dose. Typically the oil in water will comprise from 2 to 10% squalene, from 2 to 10% alpha tocopherol and from 0.3 to 3% tween 80. Preferably the ratio of squalene: alpha tocopherol is equal to or less than 1 as this provides a more stable emulsion. Span 85 may also be present at a level of 1%. In some cases it may be advantageous that the vaccines of the present invention will further contain a stabiliser.

Non-toxic oil in water emulsions preferably contain a non-toxic oil, e.g. squalene or squalene, an emulsifier, e.g. Tween 80, in an aqueous carrier. The aqueous carrier may be, for example, phosphate buffered saline.

A particularly potent adjuvant formulation involving QS21, 3D-MPL and tocopherol in an oil in water emulsion is described in WO 95/17210.

The present invention also provides a polyvalent vaccine composition comprising a vaccine formulation of the invention in combination with other antigens, in particular antigens useful for treating cancers, autoimmune diseases and related conditions. Such a polyvalent vaccine composition may include a TH-1 inducing adjuvant as hereinbefore described.

5

**Diagnostic Assays, disease monitoring, chromosomal localisation, and gene amplification**

An important aspect of the invention, in addition to the polynucleotides themselves, also relates to the use of the polynucleotides of the present invention and oligonucleotides  
10 derived from them, as diagnostic and monitoring reagents. Oligonucleotide fragments derived from the polynucleotides of the invention for use as probes or primers generally comprise at least 15 bases. Particularly preferred probes will have between 30 and 50 nucleotides. Particularly preferred primers will have between 20 and 25 nucleotides.

15 The identification of genetic or biochemical markers in blood, other biological fluids, faeces, or tissues that will enable the detection of very early changes along the carcinogenesis pathway will help in determining the best treatment for the patient and the efficacy of the treatment. Surrogate tumour markers, such as polynucleotide expression, can be used to diagnose different forms and states of cancer. The identification of expression levels of the  
20 polynucleotides of the invention will be useful in both the staging of the cancerous disorder and grading the nature of the cancerous tissue. The staging process monitors the advancement of the cancer and is determined on the presence or absence of malignant tissue in the areas biopsied. The polynucleotides of the invention can help to perfect the staging process by identifying markers for the aggressivity of a cancer, for example the presence in  
25 different areas of the body. The grading of the cancer describes how closely a tumour resembles normal tissue of its same type and is assessed by its cell morphology and other markers of differentiation. The polynucleotides of the invention can be useful in determining the tumour grade as they can help in the determination of the differentiation status of the cells of a tumour.

30

On the other hand, the polypeptide of the invention can be produced by stroma cells, in which cases, its specific expression or differential expression is a marker of disease conditions.

### Differential expression

The diagnostic assays offer a process for diagnosing or determining a susceptibility to cancers, autoimmune disease and related conditions through diagnosis by methods comprising determining from a sample derived from a subject an abnormally decreased or increased level of polypeptide or mRNA. This method of diagnosis is known as differential expression. The expression of a particular gene is compared between a diseased tissue and a normal tissue. A difference between the polynucleotide-related gene, mRNA, or protein in the two tissues is compared, for example in molecular weight, amino acid or nucleotide sequence, or relative abundance, indicates a change in the gene, or a gene which regulates it, in the tissue of the human that was suspected of being diseased.

Decreased or increased expression can be measured at the RNA level. PolyA RNA is first isolated from the two tissues and the detection of mRNA encoded by a gene corresponding to a differentially expressed polynucleotide of the invention can be detected by, for example, in situ hybridization in tissue sections, reverse transcriptase-PCR, using Northern blots containing poly A+ mRNA, or any other direct or indirect RNA detection method. An increased or decreased expression of a given RNA in a diseased tissue or surrounding tissues compared to a normal tissue or in the absence of disease suggests that the transcript and/or the expressed protein has a role in the disease. Thus detection of a higher or lower level of mRNA corresponding to SEQ ID NOS 1-6 relative to normal level is indicative of the presence of cancer in the patient.

mRNA expression levels in a sample can be determined by generation of a library of expressed sequence tags (ESTs) from the sample. The relative representation of ESTs in the library can be used to assess the relative representation of the gene transcript in the starting sample. The EST analysis of the test can then be compared to the EST analysis of a reference sample to determine the relative expression levels of the polynucleotide of interest.

Other mRNA analyses can be carried out using serial analysis of gene expression (SAGE) methodology (Velculescu et. Al. Science (1995) 270:484) , differential display methodology (For example, US 5,776,683) or hybridization analysis which relies on the specificity of nucleotide interactions.

Alternatively, where the polynucleotide sequence encodes a polypeptide, the comparison could be made at the protein level. The protein sizes in the two tissues may be compared using antibodies to detect polypeptides in Western blots of protein extracts from the two  
5 tissues. Expression levels and subcellular localization may also be detected immunologically using antibodies to the corresponding protein. Further assay techniques that can be used to determine levels of a protein, such as a polypeptide of the present invention, in a sample derived from a host are well-known to those of skill in the art. A raised or decreased level of polypeptide expression in the diseased tissue compared with the  
10 same protein expression level in the normal tissue indicates that the expressed protein may be involved in the disease.

In the assays of the present invention, the diagnosis can be determined by detection of gene product expression levels encoded by at least one sequence set forth in SEQ ID NOS: 1 - 6.  
15 A comparison of the mRNA or protein levels in a diseased versus normal tissue may also be used to follow the progression or remission of a disease.

#### Use of arrays for diagnosis

A large number of polynucleotide sequences in a sample can be assayed using  
20 polynucleotide arrays. These can be used to examine differential expression of genes and to determine gene function. For example, arrays of the polynucleotide sequences SEQ ID NOS 1 - 6 can be used to determine if any of the polynucleotides are differentially expressed between a normal and cancer cell. In one embodiment of the invention, an array of oligonucleotides probes comprising SEQ ID NOS 1 - 6 nucleotide sequence or fragments  
25 thereof can be constructed to conduct efficient screening of e.g., genetic mutations. Array technology methods are well known and have general applicability and can be used to address a variety of questions in molecular genetics including gene expression, genetic linkage, and genetic variability (see for example: M.Chee et al., Science, Vol 274, pp 610-613 (1996)).

30

"Diagnosis" as used herein includes determination of a subject's susceptibility to a disease, determination as to whether a subject presently has the disease, and also the prognosis of a subject affected by the disease.

The present invention, further relates to a diagnostic kit for performing a diagnostic assay which comprises:

- 5 (a) a polynucleotide of the present invention, preferably the nucleotide sequence of any one of SEQ ID NOs: 1 – 6, or a fragment thereof ;
- (b) a nucleotide sequence complementary to that of (a);
- (c) a polypeptide of the present invention, preferably the polypeptide encoded by a polynucleotide comprising the sequence contained in any one of SEQ ID Nos 1 – 6 or a fragment thereof; or
- 10 (d) an antibody to a polypeptide of the present invention, preferably to the polypeptide encoded by a polynucleotide comprising the sequence contained in any one of SEQ ID Nos 1 – 6.
- (e) any specific ligand to a polypeptide of the present invention

- 15 The nucleotide sequences of the present invention are also valuable for chromosomal localisation. The sequence is specifically targeted to, and can hybridize with, a particular location on an individual human chromosome. The mapping of relevant sequences to chromosomes according to the present invention is an important first step in correlating those sequences with gene associated disease. Once a sequence has been mapped to a
- 20 precise chromosomal location, the physical position of the sequence on the chromosome can be correlated with genetic map data. Such data are found in, for example, V. McKusick, Mendelian Inheritance in Man (available on-line through Johns Hopkins University Welch Medical Library). The relationship between genes and diseases that have been mapped to the same chromosomal region are then identified through linkage analysis (coinheritance of
- 25 physically adjacent genes). The differences in the cDNA or genomic sequence between affected and unaffected individuals can also be determined. In addition, gene amplification at the genomic level could be assessed and correlated with the stage of disease.

### 30 **Antibodies**

The polypeptides of the invention or their fragments or analogs thereof, or cells expressing them, can also be used as immunogens to produce antibodies immunospecific for polypeptides of the present invention. The term “immunospecific” means that the antibodies

have substantially greater affinity for the polypeptides of the invention than their affinity for other related polypeptides in the prior art.

In a further aspect the invention provides an antibody immunospecific for a polypeptide  
5 according to the invention or an immunological fragment thereof as hereinbefore defined. Preferably the antibody is a monoclonal antibody.

Antibodies generated against polypeptides of the present invention may be obtained by administering the polypeptides or epitope-bearing fragments, analogs or cells to an animal,  
10 preferably a non-human animal, using routine protocols. For preparation of monoclonal antibodies, any technique which provides antibodies produced by continuous cell line cultures can be used. Examples include the hybridoma technique (Kohler, G. and Milstein, C., Nature (1975) 256:495-497), the trioma technique, the human B-cell hybridoma technique (Kozbor *et al.*, Immunology Today (1983) 4:72) and the EBV-hybridoma  
15 technique (Cole *et al.*, Monoclonal Antibodies and Cancer Therapy, 77-96, Alan R. Liss, Inc., 1985).

Techniques for the production of single chain antibodies, such as those described in U.S. Patent No. 4,946,778, can also be adapted to produce single chain antibodies to polypeptides  
20 of this invention. Also, transgenic mice, or other organisms, including other mammals, may be used to express humanized antibodies.

The above-described antibodies may be employed to isolate or to identify clones expressing the polypeptide or to purify the polypeptides by affinity chromatography.  
25 The antibody of the invention may also be employed to prevent or treat cancer, particularly colon cancer, autoimmune disease and related conditions.

Another aspect of the invention relates to a method for inducing or modulating an immunological response in a mammal which comprises inoculating the mammal with a  
30 polypeptide of the present invention, adequate to produce antibody and/or T cell immune response to protect or ameliorate the symptoms or progression of the disease. Yet another aspect of the invention relates to a method of inducing or modulating immunological response in a mammal which comprises, delivering a polypeptide of the

present invention *via* a vector directing expression of the polynucleotide and coding for the polypeptide *in vivo* in order to induce such an immunological response to produce antibody to protect said animal from diseases.

5 It will be appreciated that the present invention therefore provides a method of diagnosing and/or treating abnormal conditions such as, for instance, cancer and autoimmune diseases, in particular, colon cancer, related to either a presence of, an excess of, or an under-expression of, any one of CASB611, CASB500, CASB501, CASB502, CASB505, or CASB507.

10

### Screening

The present invention further provides for a method of screening compounds to identify those which stimulate or which inhibit the function of any one of the CASB611, CASB500, CASB501, CASB502, CASB505, or CASB507 polypeptides. In general, agonists or  
15 antagonists may be employed for therapeutic and prophylactic purposes for such diseases as hereinbefore mentioned. Compounds may be identified from a variety of sources, for example, cells, cell-free preparations, chemical libraries, and natural product mixtures. Such agonists, antagonists or inhibitors so-identified may be natural or modified substrates, ligands, receptors, enzymes, etc., as the case may be, of the polypeptide; or may be structural  
20 or functional mimetics thereof (see Coligan *et al.*, Current Protocols in Immunology 1(2):Chapter 5 (1991)). Screening methods will be known to those skilled in the art. Further screening methods may be found in for example D. Bennett *et al.*, J Mol Recognition, 8:52-58 (1995); and K. Johanson *et al.*, J Biol Chem, 270(16):9459-9471 (1995) and references therein.

25

Thus the invention provides a method for screening to identify compounds which stimulate or which inhibit the function of the polypeptide of the invention which comprises a method selected from the group consisting of:

(a) measuring the binding of a candidate compound to the polypeptide (or to the cells or  
30 membranes bearing the polypeptide) or a fusion protein thereof by means of a label directly or indirectly associated with the candidate compound;



- (b) measuring the binding of a candidate compound to the polypeptide (or to the cells or membranes bearing the polypeptide) or a fusion protein thereof in the presense of a labeled competitor;
- (c) testing whether the candidate compound results in a signal generated by activation or inhibition of the polypeptide, using detection systems appropriate to the cells or cell membranes bearing the polypeptide;
- (d) mixing a candidate compound with a solution containing a polypeptide of claim 1, to form a mixture, measuring activity of the polypeptide in the mixture, and comparing the activity of the mixture to a standard; or
- (e) detecting the effect of a candidate compound on the production of mRNA encoding said polypeptide and said polypeptide in cells, using for instance, an ELISA assay.

The polypeptide of the invention may be used to identify membrane bound or soluble receptors, if any, through standard receptor binding techniques known in the art. Well known screening methods may also be used to identify agonists and antagonists of the polypeptide of the invention which compete with the binding of the polypeptide of the invention to its receptors, if any.

Thus, in another aspect, the present invention relates to a screening kit for identifying agonists, antagonists, ligands, receptors, substrates, enzymes, etc. for polypeptides of the present invention; or compounds which decrease or enhance the production of such polypeptides, which comprises:

- (a) a polypeptide of the present invention;
- (b) a recombinant cell expressing a polypeptide of the present invention;
- (c) a cell membrane expressing a polypeptide of the present invention; or
- (d) antibody to a polypeptide of the present invention;

which polypeptide is preferably that encoded by a polynucleotide comprising the sequence contained in any one of SEQ ID NOs:1 - 6.

It will be readily appreciated by the skilled artisan that a polypeptide of the present invention may also be used in a method for the structure-based design of an agonist, antagonist or inhibitor of the polypeptide, by:

- (a) determining in the first instance the three-dimensional structure of the polypeptide;  
(b) deducing the three-dimensional structure for the likely reactive or binding site(s) of an agonist, antagonist or inhibitor;  
(c) synthesising candidate compounds that are predicted to bind to or react with the deduced binding or reactive site; and  
(d) testing whether the candidate compounds are indeed agonists, antagonists or inhibitors.

Gene therapy may also be employed to effect the endogenous production of CASB611, CASB500, CASB501, CASB502, CASB505, or CASB507 polypeptides by the relevant cells in the subject. For an overview of gene therapy, see Chapter 20, Gene Therapy and other Molecular Genetic-based Therapeutic Approaches, (and references cited therein) in Human Molecular Genetics, T Strachan and A P Read, BIOS Scientific Publishers Ltd (1996).

15

#### **Compositions and administration**

Vaccine preparation is generally described in Pharmaceutical Biotechnology, Vol.61 Vaccine Design - the subunit and adjuvant approach, edited by Powell and Newman, Plenum Press, 1995. New Trends and Developments in Vaccines, edited by Voller et al., University Park Press, Baltimore, Maryland, U.S.A. 1978. Encapsulation within liposomes is described, for example, by Fullerton, U.S. Patent 4,235,877. Conjugation of proteins to macromolecules is disclosed, for example, by Likhite, U.S. Patent 4,372,945 and by Armor et al., U.S. Patent 4,474,757.

The amount of protein in each vaccine dose is selected as an amount which induces an immunoprotective response without significant, adverse side effects in typical vaccinees. Such amount will vary depending upon which specific immunogen is employed. Generally, it is expected that each dose will comprise 1-1000µg of protein, preferably 2-100µg, most preferably 4-40µg. An optimal amount for a particular vaccine can be ascertained by standard studies involving observation of antibody titres and other responses in subjects. Following an initial vaccination, subjects may receive a boost in about 4 weeks.

**Definitions**

“Isolated” means altered “by the hand of man” from the natural state. If an “isolated” composition or substance occurs in nature, it has been changed or removed from its original environment, or both. For example, a polynucleotide or a polypeptide naturally present in a living animal is not “isolated,” but the same polynucleotide or polypeptide separated from the coexisting materials of its natural state is “isolated”, as the term is employed herein.

“Polynucleotide” generally refers to any polyribonucleotide or polydeoxribonucleotide, which may be unmodified RNA or DNA or modified RNA or DNA including single and double stranded regions.

“Variant” refers to a polynucleotide or polypeptide that differs from a reference polynucleotide or polypeptide, but retains essential properties. A typical variant of a polynucleotide differs in nucleotide sequence from another, reference polynucleotide. Changes in the nucleotide sequence of the variant may or may not alter the amino acid sequence of a polypeptide encoded by the reference polynucleotide. Nucleotide changes may result in amino acid substitutions, additions, deletions, fusions and truncations in the polypeptide encoded by the reference sequence, as discussed below. A typical variant of a polypeptide differs in amino acid sequence from another, reference polypeptide. Generally, differences are limited so that the sequences of the reference polypeptide and the variant are closely similar overall and, in many regions, identical. A variant and reference polypeptide may differ in amino acid sequence by one or more substitutions, additions, deletions in any combination. A substituted or inserted amino acid residue may or may not be one encoded by the genetic code. A variant of a polynucleotide or polypeptide may be a naturally occurring such as an allelic variant, or it may be a variant that is not known to occur naturally. Non-naturally occurring variants of polynucleotides and polypeptides may be made by mutagenesis techniques or by direct synthesis.

“Identity,” as known in the art, is a relationship between two or more polypeptide sequences or two or more polynucleotide sequences, as determined by comparing the sequences. In the art, “identity” also means the degree of sequence relatedness between polypeptide or polynucleotide sequences, as the case may be, as determined by the match between

- strings of such sequences. "Identity" and "similarity" can be readily calculated by known methods, including but not limited to those described in (Computational Molecular Biology, Lesk, A.M., ed., Oxford University Press, New York, 1988; Biocomputing: Informatics and Genome Projects, Smith, D.W., ed., Academic Press, New York, 1993;
- 5 Computer Analysis of Sequence Data, Part I, Griffin, A.M., and Griffin, H.G., eds., Humana Press, New Jersey, 1994; Sequence Analysis in Molecular Biology, von Heinje, G., Academic Press, 1987; and Sequence Analysis Primer, Gribskov, M. and Devereux, J., eds., M Stockton Press, New York, 1991; and Carillo, H., and Lipman, D., *SIAM J. Applied Math.*, 48: 1073 (1988). Preferred methods to determine identity are designed to
- 10 give the largest match between the sequences tested. Methods to determine identity and similarity are codified in publicly available computer programs. Preferred computer program methods to determine identity and similarity between two sequences include, but are not limited to, the GCG program package (Devereux, J., et al., *Nucleic Acids Research* 12(1): 387 (1984)), BLASTP, BLASTN, and FASTA (Atschul, S.F. et al., *J. Molec. Biol.* 215: 403-410 (1990). The BLAST X program is publicly available from
- 15 NCBI and other sources (BLAST Manual, Altschul, S., et al., NCBI NLM NIH Bethesda, MD 20894; Altschul, S., et al., *J. Mol. Biol.* 215: 403-410 (1990). The well known Smith Waterman algorithm may also be used to determine identity.
- 20 The preferred algorithm used is FASTA. The preferred parameters for polypeptide or polynucleotide sequence comparison using this algorithm include the following:  
Gap Penalty:12  
Gap extension penalty: 4  
Word size: 2, max 6
- 25 Preferred parameters for polypeptide sequence comparison with other methods include the following:  
1) Algorithm: Needleman and Wunsch, *J. Mol Biol.* 48: 443-453 (1970)  
Comparison matrix: BLOSSUM62 from Hentikoff and Hentikoff, *Proc. Natl. Acad. Sci. USA.* 89:10915-10919 (1992)
- 30 Gap Penalty: 12  
Gap Length Penalty: 4

A program useful with these parameters is publicly available as the "gap" program from Genetics Computer Group, Madison WI. The aforementioned parameters are the default parameters for polypeptide comparisons (along with no penalty for end gaps).

5 Preferred parameters for polynucleotide comparison include the following:

1) Algorithm: Needleman and Wunsch, J. Mol Biol. 48: 443-453 (1970)

Comparison matrix: matches = +10, mismatch = 0

Gap Penalty: 50

Gap Length Penalty: 3

10

A program useful with these parameters is publicly available as the "gap" program from Genetics Computer Group, Madison WI. The aforementioned parameters are the default parameters for polynucleotide comparisons.

15 By way of example, a polynucleotide sequence of the present invention may be identical to the reference sequence of SEQ ID NO:1, that is be 100% identical, or it may include up to a certain integer number of nucleotide alterations as compared to the reference sequence. Such alterations are selected from the group consisting of at least one nucleotide deletion, substitution, including transition and transversion, or insertion, and  
 20 wherein said alterations may occur at the 5' or 3' terminal positions of the reference nucleotide sequence or anywhere between those terminal positions, interspersed either individually among the nucleotides in the reference sequence or in one or more contiguous groups within the reference sequence. The number of nucleotide alterations is determined by multiplying the total number of nucleotides in SEQ ID NO:1 by the  
 25 numerical percent of the respective percent identity (divided by 100) and subtracting that product from said total number of nucleotides in SEQ ID NO:1, or:

$$n_n \leq x_n - (x_n \cdot y),$$

wherein  $n_n$  is the number of nucleotide alterations,  $x_n$  is the total number of nucleotides in SEQ ID NO:1, and  $y$  is, for instance, 0.70 for 70%, 0.80 for 80%, 0.85 for 85%, 0.90  
 30 for 90%, 0.95 for 95%, etc., and wherein any non-integer product of  $x_n$  and  $y$  is rounded down to the nearest integer prior to subtracting it from  $x_n$ . Alterations of a polynucleotide sequence encoding the polypeptide of SEQ ID NO:2 may create nonsense,

missense or frameshift mutations in this coding sequence and thereby alter the polypeptide encoded by the polynucleotide following such alterations.

“Homolog” is a generic term used in the art to indicate a polynucleotide or polypeptide sequence possessing a high degree of sequence relatedness to a subject sequence. Such relatedness may be quantified by determining the degree of identity and/or similarity between the sequences being compared as hereinbefore described. Falling within this generic term are the terms “ortholog”, meaning a polynucleotide or polypeptide that is the functional equivalent of a polynucleotide or polypeptide in another species and “paralog” meaning a functionally similar sequence when considered within the same species.

## **EXAMPLES**

### **Example 1**

#### **Subtractive cDNA cloning of colon tumour-associated antigen (TAA) candidates.**

5           Subtractive cDNA libraries are produced using standard technologies. Briefly, total RNA is extracted from frozen (-70°C) colon tumour and matched normal colon samples using the TriPure reagent and protocol (Boehringer). Target RNA is prepared by pooling total RNA from three tumour samples (30 µg each). Driver RNA is prepared by pooling  
10           total RNA from three matched normal colon samples (10 µg each) and total RNA from seven normal tissues other than colon(brain, heart, kidney, liver, bladder, skin, spleen; 10 µg each). Total RNA from non-colon normal tissues is purchased from InVitrogen.

          Messenger RNA is purified from total RNA using oligo-dT magnetic bead technology (Dynal) and quantified by spectrofluorimetry (BioRad).

15           Target and driver mRNA are reverse transcribed into cDNA using one of two strategies: 1) Target sequences for PCR oligonucleotides are introduced onto the ends of the newly synthesised cDNA during reverse transcription using the template switching capability of reverse transcriptase (ClonTech SMART PCR cDNA synthesis kit). 2) Target  
20           and driver mRNA are reverse transcribed into cDNA using an oligo-dT primer and converted to double-strand cDNA; the cDNA is cleaved with RsaI and linkers for PCR amplification are ligated onto the extremities of the cDNA fragments.

          In both cases, target and driver cDNA are amplified by long range PCR (ClonTech  
25           SMART PCR Synthesis Kit and Advantage PCR Polymerase Mix) and used as starting material for subtractive cloning. For amplification, cycling conditions and optimisation of the number of PCR cycles are as described in the Advantage PCR protocol.

          Two subtractive cloning strategies are used: ClonTech PCR SELECT (see  
30           ClonTech kit protocol and N. Gurskaya et al. 1996. Analytical Biochemistry: 240, 90) and cRDA(M. Hubank and D. Schatz. 1994. Nucleic Acids Research: 22, 5640) . When the PCR SELECT protocol is used, the primary PCR SELECT subtraction products are submitted to

a supplementary round of cRDA subtraction. When the cRDA protocol is used, two consecutive cycles of cRDA subtraction are performed.

In each case the products of both cycles of subtraction are cloned into pCR-TOPO (Invitrogen) and transformed into E. coli to produce a subtracted cDNA plasmid library.

An alternative strategy is also followed: subtraction of normal colon sequences and sequences from non-colon normal tissues are subtracted in separate hybridisations. In this case, target and driver RNA are assembled for the first subtraction as above with the exception that non-colon RNA is left out of the driver pool and amounts of normal colon are increased to 10 µg. Preparation of target and driver cDNA and subtractive hybridisation are performed as described above. A second subtraction is then performed on the products of the first subtraction, but the driver is now composed of a pool of normal colon and normal non-colon mRNA from the seven normal tissues.

15

#### **Differential Screening of cDNA arrays.**

Identification of tumour-associated genes in the subtracted cDNA library is accomplished by differential screening.

Total bacterial DNA is extracted from 100 µl over-night cultures. Bacteria are lysed with guanidium isothiocyanate and the bacterial DNA is affinity purified using magnetic glass (Boehringer). Plasmid inserts are recovered from the bacterial DNA by Advantage PCR amplification (Clontech). The PCR products are dotted onto two nylon membranes to produce high density cDNA arrays using the Biomek 96 HDRT tool (Beckman). The spotted cDNA is covalently linked to the membrane by UV irradiation. The first membrane is hybridised with a mixed cDNA probe prepared from the tumour of a single patient. The second membrane is hybridised with an equivalent amount of mixed cDNA probe prepared from normal colon of the same patient. The probe cDNA is prepared by PCR amplification as described above and is labelled using the AlkPhos Direct System (Amersham).

Hybridisation conditions and stringency washes are as described in the AlkPhos Direct kit. Hybridised probe is detected by chemiluminescence. Hybridisation intensities for each cDNA fragment on both blots (see figure 1) are measured by film densitometry or direct measurement (BioRad Fluor-S Max). The ratio of the tumour to normal hybridisation

25  
30



intensities (T/N) is calculated for each gene to evaluate the degree of over-expression in the tumour. Genes which are significantly over-expressed in colon tumours are followed-up. Significance is arbitrarily defined as one standard deviation of the T/N frequency distribution. Differential screening experiments are repeated using RNA from multiple patient donors (>18) to estimate the frequency of over-expressing tumours in the patient population.

In addition, the DNA arrays are hybridised with mixed cDNA probes from normal tissues other than colon (see list above) to determine the level of expression of the candidate gene in these tissues.

## **Example 2**

### **Real-time RT-PCR analysis**

Real-time RT-PCR (U. Gibson. 1996. Genome Research: 6,996) is used to compare mRNA transcript abundance of the candidate antigen in tumour and normal colon tissues from multiple patients. In addition, mRNA levels of the candidate gene are re-evaluated by this approach in a panel of normal tissues.

Total RNA is extracted from snap frozen colon tissue biopsies using TriPure reagent (Boehringer). Total RNA from normal tissues is from InVitrogen as above. Poly-A<sup>+</sup> mRNA is purified from total RNA after DNAase treatment using oligo-dT magnetic beads (Dyna). Quantification of the mRNA is performed by spectrofluorimetry (BioRad) using SybrII dye (Molecular Probes). Primers for amplification are designed with the Perkin-Elmer Primer Express software using default options for TaqMan amplification conditions.

Real-time reactions are assembled according to standard PCR protocols using 2 ng of purified mRNA for each reaction. SybrI dye (Molecular Probes) is added at a final dilution of 1/75000 for real-time detection. Amplification (40 cycles) and real-time detection is performed in a PE 7700 system. Ct values are calculated using the 7700 Sequence Detector software for the tumour (CtT) and normal (CtN) samples of each patient. The difference between Ct values (CtN-CtT) is a direct measure of the difference in transcript levels between the tumour and normal tissues. As Ct values are log-linearly

related to copy number and that the efficiency of PCR amplification under the prevailing experimental conditions is close to the theoretical amplification efficiency,  $2^{(C_{IN}-C_{IT})}$  is an estimate of the relative transcript levels in the two tissues (i.e. fold mRNA over-expression in tumor). The percentage of over-expressing patients and the average level of mRNA over-expression in the tumours of these patients is calculated from the data set of patients.

The values indicated in table 1 are "Colon range Ct" and "colon mean Ct", representing the range and mean of Ct values for colon sample pairs, to assess the global mRNA abundance, 40 corresponding to the maximum Ct and thus to an absence of amplification.

TABLE 1: Expression of CASB genes in colon cancer and normal colon

| Gene name | relevant /tested patients | colon range Ct | colon mean Ct |
|-----------|---------------------------|----------------|---------------|
| CASB611   | 15                        | 20-40          | 28.4          |
| CASB500   | 18                        | 17-26.5        | 22            |
| CASB501   | 17                        | 20.5-40        | 25.7          |
| CASB502   | 18                        | 31.4-37.3      | 25.6          |
| CASB505   | 17                        | 22.6-40        | 29            |
| CASB507   | 9                         | 29.5-40        | 33.8          |

TABLE 2:

| Gene name | Patients over-expressing CASB in colon tumours (%) | Average level of over-expression in colon tumours (fold) |
|-----------|--|--|
| CASB500   | 94%  | 5  |
| CASB501   | 71%  | 5  |
| CASB502   | 61%  | 7  |
| CASB505   | 76%  | 18   |
| CASB507   | 78%  | 7  |

In addition, Ct values for 12 normal tissues are reported. The tissues tested are bladder, brain, breast, cervix, heart, kidney, liver, lung, oesophagus, placenta and uterus. The range and mean of the Ct value is an indicator of the specificity of tissue expression. Thus, 2 types of sequences are described, classified into categories A and B. A sequence

of category A has a colon-specific expression. A sequence of category B is overexpressed in colon cancer.

5 TABLE 3: Expression of CASB genes in normal tissues

| Gene name | Normal<br>Tissues | NT<br>Range Ct | NT<br>mean<br>Ct | class | Comments            |
|-----------|-------------------|----------------|------------------|-------|---------------------|
| CASB611   | 12                | 22-40          | 33               | A     | High only in rectum |
| CASB500   | 12                | 19-26.5        | 22.5             | B     |                     |
| CASB501   | 12                | 22.5-28.5      | 25.3             | B     | High in lung        |
| CASB502   | 12                | 21.5-28.5      | 23               | B     |                     |
| CASB505   | 12                | 14-35          | 26.5             | B     |                     |
| CASB507   | 12                | 33.5-38.5      | 35               | B     |                     |

NT range Ct and mean Ct: Ct values (range and mean) in normal tissues

10

### **Example 3**

#### **Identification of the full length cDNA sequence**

- Colon tumour cDNA libraries are constructed using the Lambda Zap system (Stratagene)
- 15 from 2 µg of polyA<sup>+</sup> mRNA as described in the supplied protocol.  $1.5 \times 10^6$  independent phage are plated for each screening of the library. Phage plaques are transferred onto nylon filters, hybridised using a cDNA probe labelled with AlkPhos Direct (Amersham Pharmacia) and positive phage are detected by chemiluminescence. The positive phage are excised from the agar plat, eluted in 500µl SM buffer and confirmed by gene-specific PCR.
- 20 Eluted phage are converted to single strand M13 bacteriophage by *in vivo* excision. The bacteriophage is then converted to double strand plasmid DNA by infection of *E. coli*. Infected bacteria are plated and submitted to a second round of screening with the cDNA probe. Plasmid DNA is purified from positive bacterial clones and submitted to Southern blot analysis to estimated the size of the cDNA inserts. CDNA inserts from multiple
- 25 independent clones are sequenced on both strands.

- When the full length gene cannot be obtained directly from the cDNA library, missing sequence is isolated using RACE technology (Marathon Kit, ClonTech.). This approach relies on reverse transcribing mRNA into double strand cDNA, ligating linkers onto the ends
- 30 of the cDNA and amplifying the desired extremity of the cDNA using a gene-specific

primer and one of the linker oligonucleotides. Marathon PCR products are cloned into a plasmid and sequenced.

- 5 All publications, including but not limited to patents and patent applications, cited in this specification are herein incorporated by reference as if each individual publication were specifically and individually indicated to be incorporated by reference herein as though fully set forth.

**SEQUENCE INFORMATION****SEQ ID 1:**

CTGCATTGAGTTTATTTATTTATTTACTTGTGTTGGATGCTCTTCCAGCAGACTCATGTCAGGAG  
5 GCTAGGAAGGATTCCCTTGGGGACCACTGGATGCTGGTAGTTAAATGCCAGGAGGCTGAATGGACCTGAAGA  
TGGAGGAGACTCTGCAGTCTTGGTCAGCCACCCTTGGGTGCTTGCCACCCTGCACTCTAGCAGGATTGATGG  
TCTCTGGATTGTAGCTGTGACCGTCATGGTGAATGCTCGGTGGTTTGCAGTGGAGAGGCCACATGGTG  
GCGACTGAGGCCCTGTGGGTGAGGTTGGCTCATGATAGCTCTGAAAGTTGATGGCACAATTGAGACAAGGA  
CGTGGAGTTCTGGAACTTTCCAGGTCTTGCAGCCCGAGATGAAGCCCTCCAACATTGCAGCTCACTGCT  
10 GTGAAGGCTGAAAGGGGAGACAGGAGGCTGCTGGGTACCTCAGGTGCCAGGCATGAGGTTCTCCGCTGTCTG  
GAGTCTTGCTTTGGGAGTGCTACACCTGTGGGACTCCCCAGGCACCACCCAAAGCGGGGCTCTGGCTCAAG  
CTTGTCATGGTTTACAGAGCCTCACATGATGTCCTTTCAGGTTTGTGAGTTGGGGCTAGGGAACGCTGGCA  
GCAGAGCCTGGTTCTCTGCCTGACCAGGCCGTTGGCAGGACGCCTCTTCCCCTCTGTGGAGAAGATGGAGA  
AGCAGAGAAGCAGGATACAGAGCAGGCTGGAAGGACTAGAAGCCTCATGGTGAGAAGGCAGATTTCCCTC  
15 GTTCCGAAT

**SEQ ID NO 2:**

TTTTTTAGGGTTCAGTTCAGCTGATTTTATTTCTTCTCAAAAAAGTTATTTACAGAAGGTATATATCAA  
CAATCTGACAGGCAGTGAACCTTGACATGATTAGCTGGCATGATTTTTCTTTTTTCCCCAACATTGTT  
20 TTTGTGGCCTTGAATTTTAAGACAAATATTCTACACGGCATATTGCACAGGATGGATGGCAAAAAAGTTT  
AAAAACAAAAACCCTTAACGGAACGCTTAAAAAGGCAGACGTTTCTAGTGCCTGTCATGTTATATTAAACA  
TACATACACACAATCTTTTGGCTTATTATAATACAGACTTAAATGTACAAAGATGTTTCCACTTTTTTCAA  
TTTTTAAACACAACAGCTATAAACCTGAACACATATGCTATCATCATGCCATAAGACTAAACAATTATATT  
TAGCGACAAGTAGAAAGGATTAAATAGTCAAATACAAGAATGAAAAACGCAGTACATAGTGTGCGGAACCTCA  
25 AATCGGCATTTAGATAGATCCAGTGGTTTAAACGGCACGTTTTTGCTTATAAAAAAGTGCAAAAAAGATGT  
GGTTTACAAGTTAAAGCTACAGAATCCCTTTTGGCTGTAATTGCACCAAGTTTAAAGCCTCTGGACAGAGCA  
GTATTTTCGTTTAAACCTTTGTTTTCTTAAAGCTTACAGTGTGTTGGCTAATTCTCCTCCCTTTTACAAG  
ACGGGGGCCGGAGGGTGGACACTGGTGGCAGGTTAAGGGATACTGTCACTTTAAAGCCTGCAGATTGAAG  
TGTAACATGGAGAAATTAGGGGCTGATTTTTTAACTGTGTGAGATATTAACCAGCCGCCCTGTTATAAAA  
30 TCAGGAAATCCAAACAGCGATTTACACCGATTAACACCCCTTTATATATTTTTTACAAAAATACACTGAGA  
AAATAATCAAACGTTTTTCATCTCTCTGCTTTTTTTGTTTTTAAAAAGTGTCAAAAGCTACATTTAAATA  
TAAAAAATTAAAGTTAAACTCTAGCCCTTCAGTGAAGGAGACGTAAATGGCGTGGGTAAACAACACTAC  
CAAAAAAAGAAAAAAGAAAAAAGGAAAGGAAGGAATAAAGAAATAAAGGAAGTAAAAAGAAAG  
GAAAGAAAAAAGGGACAAAAGAAAAAATATGTTGGCCAGTATAAATACGTCCACATATAAAATGGCAT  
35 CTGATTACATTTACAAGGAAAAAAGAAATACAGGATGGAGCATCGGTGAAGGAAAAAACACGCTTCTCAT  
TTACACCTATAAGGAATAAACACACACACTGAGAAAAAATTTGGCCCTGAATTGTTTTTTAAAGTCCAGCA  
CAGATTTGAGTTGCGTTTGAATCCTTTAAAGAGTTAAGAATGAAAAAAGCTGGTGATAATTTGTTGTAGG  
AATCAACATAGCGCCATCTATCTGCTTTTATATTATCTACACTATTTTAAAACTGCTCAACAGTCTTA

TACAGAAATCTTTAAAAGATAGACAGGATAACATGCTATATTAACCCACCATTGAAATAATCCAACACCAT  
CACGATTCCGATTAAGAGAAGNAAAAAAT

**SEQ ID NO: 3**

5 GAGAAATTTTAATTTGGTGTACAAAACCTTTTAGTAGTTAGTGAATGAGTAAGATGATTGTGTATTGTGCAG  
TTTGGCTTTTAAGAGCAAAGGAGAACTTAGAGACTATGGGTTCTAGATGCTGTGTCTGCCTTTCACAGTTTCT  
CTGAACCTTTGTGATTGAACCTGGATTCATATTGGAGTTATATATCAGACAAAACCAAAGTCATTGGGCACAA  
GCAGAATAAGTCCAAATTTAAAGTATAATCATGAGTTTGTAAATAATGCTTTATATACTTCTTTAAAATAT  
ACTTATTTAATATTAAGGATTTTTTTTTTCCCTGAATCCATATAATTTGAAAGGTCTAAGGGGAAAACTTGT  
10 TTTGAGCCACTTTATGATGTGAATTCACCTTCATAGGGAAATAACATGCTTGGTTTTAATGTACTTTAACCC  
TTTTAAAGTACACTGGACCAAATTTTAGAGAGCATATACATAGTGTATAAAAT

**SEQ ID NO: 4**

15 CAGTCTCTGAAATCTGAAATCCNGTTTCTGAAATCTTGAATTGCCTCCCANTTGGCCATTTTCCCTTTGAGC  
CAAAAGTTTCAGATTTTGGAGCATTTTGGATTGGGATGCTCAACCTGTGTAAATATTTGTTAATATATAA  
TGTAATAATTTTAAAGATCTCAAACAATTAATTTGAAAATTACCATTTAATTATGCCAGCATTTGGTCATT  
CTCCCTAATTTTCTAGTCCGCTCTAATCTGTCTCCTCAGTAATAATAAAAGATAGCACATTCATTTATT  
ATAAAGCCCCAAGTAGGTATAAGTTTAACTCATAACTGTTGAAGAGTCAGCTTTATGTCTGAAATAATCAC  
20 TTTATACAATGT  
TATACCTGGGGTAGTGCAGCATGCACATTTCAAGATTGTAAAGACTACTACCCGAATCCAAAATTGAGAGAG  
AACCTGTTTATATAAACTCATTAGTGAACACTCTTCTTAACAGCTACAGAGGCTGCCAGTAGCTATGATGT  
TAATTCTCTACTTTCATGCATCTGTAGGCTAAATGGAAAATTAAGATATGACTGAAGAGAAGCTCCTCTTCAT  
ATTCAGGTCTGCAACTTTTCTCTTTAAAAGGAGAGAAATTAAGATCAATGAATACCAATAGTGGAGAACTTT  
25 ACAACGTATGAATTAACCTCTACAGTCAAGGGGTCACACTTATAACGTTGTATCAGTAAGTCACCTAACGA  
TCTACTTGTATCATGTCTAGCAAACTTCAGTGTCCCCATCTTACTTGTCTTCTCCATGACACCTGAATTTTC  
CTGATGCTTTTCTTTCTCTTGACTTCAATGACATCTCCCATGTTTACCATTTTTCCCTATTTCTATGGTTC  
CCTTTTCAGTCTTCCAGATGCTGTGCTCTCCTTCTTAACTAACCTGTCAGATGGCAATATTTCTCATAG  
TTCTGACTTAGGCACCTTGCTTCTGTAAACATACCTTGTGTGAAACACTCAAAGCTT

30

**SEQ ID NO: 5**

ACATTTTTTATTACGTAAACAGTAAAAACATGAAACCTATCAAGTTTGCTAAAACCAAATATTGCAGGCTAAG  
ACAAGACTTTTCAGATTGGCTATCATTTAATAGAATAAGAAAGACCTCCTTAATTCTGTTTCCCTTTCTTA  
35 AGGGTAAGTATCATCTGCTCATTAACCTGACAAATGGTGTGGTCAAGTAGCAAGGCTGCAGCCAAAAGACCT  
ACACTGGATATCTGACTCCTTATCTTTTCAATGTTGTGTGACCTTGAGCAAGTTACTTTAACCTTTTAAAGCCT  
CTGNTTTTTTCTTCTATAACAGAGACAGTAATATTTATCACTGGTTTTTTTGNAGAGATTAATGGGCGTAT  
CTTATGTAATCACCTAAAATATTTTCTAGGTAGTAGGCATCAAATAAGCCTGCTATTACTACTTTTGGGA  
GAAGATAACCAGGTCCTTGACATTACCACCAAAGGTACCCTG

40

**SEQ ID NO: 6**

CCGACGTCGACTATCCAAGATGTACTATTATTGGCATATAGTTTCAAAAATTCACAGAAGGGAGCCAGGTGC  
TCTCATTGCCTAACAAAATAATGAAATATGTATTCTTAAACATCTTGACATACAGTATAAAGGGCACTC  
45 AGCAAGTGCTTTTGTAGTACTGATTTTAAATGAGTAGATTACAGGAATACCAGAAAACGCACACACACACA  
GATAATTGGTTGAAAATACGAGGTTAGAACTTAACCATTTGAAGAGGAGTAAGGCTACTTAAAGCGTTAAAA  
ACTAATTGGGTAAGGTATGGTTGACCCAGCTACTTCATTTGCCTTACGATGTATATTCAATTAACCTAAGTCA  
CTCACCTTCTCTGTGATGTTGATGGTTTGGTATAGTAAACATGAGATATGATTAAAGGTGATTACAGGGATAG  
ATCAAGTGTCTGCCTAAGTAAATCTGGGNTTCAATTTTTTCTAGACATCTTGA

50

**SEQ ID NO: 7**

TTGAGTTTATTATTATTATTACTTGTGTGTCTGGATGCTCTTTCCAGCAGACTCATGTCCAGGAGGCTA  
GGAAGGGATTCTTGGGGACCACTGGATGCTGGTAGTTAAATGCCAGGAGGCTGAATGGACCTGAAGATGGA  
55 GGAGACTCTGCAGTCTTGGTCAGCCACCTTGGGTGCTTGCCACCCTGCACCTAGCAGGATTGATGGTCTC  
TGGATTTGTAGCTGTGACCGGTCATGGTGAATGCTCGGTGGTTGCACTGGAGAGGCCACATGGTGGCGA

CTGAGGCCCTGTGGGGTGAGGTTGGGCTCATGATAGCTCTGAAAGTTGATGGCACAATTGAGACAAGGACGTG  
GAGTTCTGGAACCTTTCCAGGGTCCCTGCAGCCCAGATGAAGCCCTCCAACATTGCAGCTCACTGCTGTGA  
AGGCTGAAAGGGGAGACAGGAGGCTGCTGGGTACCTCAGGTGCCAGGCATGAGGTTCTCCGCTGTCTGGAGT  
CTTGCTTTGGGAGTGCTACACCTGTGGGACTCCCCAGGCACCACCCAAAGGCGGGGCTCTGGCTCAAGCTTG  
5 CATGGTTTACAGAGCCTCACATGATGTCCTTTCAGGTTTGTGAGTTGGGGCTAGGGACTCGGTGGCAGCAG  
AGCCTGGTTCTCTGCTGACAGGCCGTTGGCAGGACGCCTCTTCCCCTCTGTGGAGAAGATGGAGAAGCA  
GAGAAGCAGGATACAGAGCAGGCTGGAAAGGACTAGAAGCCTCATGGTGAGAAGGCAGATTTTCCCTCGTTC  
CGAAT

10 **SEQ ID NO: 8**

TTTTTTAGGGTTCAGTTCAGCTGATTTTATTTCCCTTCTCAAAAAAGTTATTTACAGAA  
GGTATATATCAACAATCTGACAGGCAGTGAACCTTGACATGATTAGCTGGCATGATTTTTT  
CTTTTTTTTCCCCAAACATTGTTTTTGTGGCCTTGAATTTTAAAGACAAATATTCTACAC  
15 GGCATATTGCACAGGATGGATGGCAAAAAAAGTTTAAAAACAAAAACCTTAAACGGAAC  
TGCCCTTAAAAAGGCAGACGTCCTAGTGCCTGTCTATGTTATATTAAACATACATACACACA  
ATCTTTTTTGTCTTATTATAATACAGACTTAAATGTACAAAGATGTTTTCCACTTTTTTCAA  
TTTTTAAACACAACAGCTATAAACCTGAACACATATGCTATCATCATGCCATAAGACTAA  
AACAATTATATTTAGCGACAAGTAGAAAGGATTAAATAGTCAAATACAAGAATGAAAAAC  
20 GCAGTACATAGTGTGCGGAACCTCAATCGGCATTAGATAGATCCAGTGGTTTAAACGGC  
ACGTTTTTGTCTTATAAAAAAAGTGCAAAAAAGATGTGGTTTACAAAGTTAAAGCTACAGAA  
TCCCTTTTTGTCTGTAATTGCACCACTTTTAAAGCCTCTGGACAGAGCAGTATTTCTGTTTA  
AACTTTGTTTTCTTAAAGCTTACAGTGTGTTGGCTAATTCTCTCCCTTTTTTACAAAG  
ACGGGGGCGGAGGGTGGACACTGGTGGCAGGTTAAGGGTACTGTCACTTTAAGAAGCC  
25 TGCAGATTGAAGTGTAAACATGGAGAAATTAGGGGCTGATTTTTTAAAGTGTGTGAGATA  
TTAACCAGCCGCCCTGTTATAAAATCAGGAAATCCAAACAGCGATTTACACCGATTAAACA  
CCCCCTTTATATATTTTTTACAAAAATACACTGAGAAAAATCAACGTTTTTCTCTCT  
CTTGCTTTTTTTTTGTTTTTAAAAAGTGTCAAAAGTCTACATTTAAATATAAAAAATTTAA  
AGTTAAACTCTAGCCCTTCAGTGAAGGAGACGTAAAATGGCGTGGGTAACAACAACACTAC  
30 CAAAAAAGAAAGAAAAAAGAAAAAAGGAAAGGAAGGAATAAGAAATAAAGGAA  
GTAAAAAGAAAGGAAAGAAAAAAGGACAAAAGAAAAATATGTTTGGCCCACTATAAA  
TACGTCCACATATAAAATGGCATCTGATTACATTTACAGGAAAAAGAAATACGAGGAT  
GGAGCATCGGTGAAGGAAAAAACACGCTCTCTCATTACACCTATAAGGAATAAACACA  
CACACTGAGAAAAAATTTGGCCCTGAATTGTTTTTTAAAGTCCAGCACAGATTTGAGTT  
35 GCGTTTGAATCCTTTAAAGAGTTAAGAATGAAAAAAGCTGGTGATAATTTCTGTTGAGG  
AATCAACATAGCGCCATCTATCTGCTTTTATATTATCTTACACTATTTTAAAAACTGC  
TCAACAGTCTTATACAGAAATCTTTAAAGATAGACAGGATAACATGCTATATTAACCCC  
ACCATTGAAATAATCCAACCCATCACGATTCGGATTAAGAGAAGNAAAAA

40 **SEQ ID NO: 9**

GAGAAATTTTAAATTTGGTGTACAAAACCTTTAGTAGTTAGTGAATGAGTAAGATGATTTGTGTATTGTGCAG  
TTTGGCTTTAAGAGCAAAGGAGAACTTAGAGACTATGGGTTCTAGATGCTGTGCTGCCTTTACAGTTTCT  
45 CTGAACCTTTGTGATTGAACCTGGATTCAATTTGGAGTTATATATCAGACAAAAACCAAGTCATGGGCACAA  
GCAGAATAAGTCCAAATTAAAGTATAATCATGAGTTTGTAAATAATGCTTTATATACTCTTTAAATAT  
ACTTATTTAATATTAAGGATTTTTTTTCCCTGAATCCATATAATTTGAAAGGTCTAAGGGGAAAAACTTGT  
TTTGAGCCACTTTATGATGTGAATTCATTCATAGGGAAATAACATGCTTGGTTTTTAAATGTACTTTAACC  
TTTTAAAGTACACTGGACCAATTTTAGAGAGCATATACATAGTGTATAAAAT

50 **SEQ ID NO: 10**

TTTTTTTTTTTAAAGCTTTTGGTGTTCACACAAGGTATGTTACAGAAGACAAGTGCCTAAGTCAGAACTAT  
GAGAAATATTGCCATCTGACAGGTTAGATTAAAGAAGGAGAGCACAGCATCTGGGAAGACTGAAAAGGGAAC  
CATAGAAATAGGGGAAAAATGGTGAACATGGGAGATGTCATTGAAGTCAAGAGAAAGAAAGCATCAGGAAA  
55 ATTCAGTGTCATGGAGAAAGCAAGTAAGATGGGGACAGTGAAGTTTGCTAGACATGATACAAGTAGATCGT  
TAGGTGACTTACTGATACAACGTTATAAGTAGTGACCCCTTAAGTGTAGGAGTTAATTCATACGTTGTAAAG  
TTCTCCAGTATTGGTATTCATTGATCTTAATCTCTCTCTTTTANAGAGNAAAGTTGCAGACGTGAATATGA  
GGAGG

**SEQ ID NO: 11**

CTAATACGACTCACTATAGGGCAGCGTGGTCGCGGCCGAGGTACATTTTTATTACGTAACAGTAAAAACAT  
GAAACCTATCAAGTTTGCTAAAACCAAATATTGCAGGCTAAGACAAGACTTTCAGATTGGCCTATCATTTAA  
5 TAGAATAAGAAAAGACCTCCTTAATTCTGTTTCCCTTTCTTAAGGGGTAAGTATCATCTGCTCATTAAGTGA  
CAAATGGTGTGGTCAAGTAGCAAGGCTGCAGCCAAAAGACCTACACTGGATATCTGACTCCTTATCTTTTCAT  
GTTGTGTGACCTTGAGCAAGTTACTTTAACCTTTTAAAGCCTCTGTTTTTTCTTCTATAACAGAGACAGTA  
ATATTTATCACTGTTTTTTTTTTGTAGAGATTAAATGGCGTATCTTATGTAAATCACCTAAAATATTTTTCTA  
10 GTAGTAGGGCATCAAAATAAGCCCTGCTATTACTACTTTTTGAAGAAGATAAACCAAGTCCCTGGAACATTA  
ACCACCAAAAAGTACCACTGGTTTTACACCCCCCCC

**SEQ ID NO: 12**

CCTATGCATCAAGCTTGGTACCGAGCTCGGATCCACTAGTAACGGCCGCCAGTGTGCTGGAATTCGCCCTTA  
15 CCGACGTCGACTATCCAAGATGTACTATTATTGGCATATAGTTTCAAAAATTCACAGAAGGGAGCCAGGTGC  
TCTCATTTGCCAACAATAATGGAAATATGTATTCTTCTAACATCTTGACATACAGTATAAAGGGCACTC  
AGCAAGTGCTTTTTAGTTAGACTGATTTTAAATGAGTAGATTACAGGAATACCAGAAAACGCACACACACA  
GATAATTGGTTGAAAATACGAGGTTAGAACTTAACCATTGAAGAGGAGTAAGGCTACTTAAAGCGTTAAAA  
ACTAATTGGGTAAGGTATGGTTGACCCAGCTACTTCATTTGCCTTACGATGTATATTCAATTAAGTCA  
20 CTCACCTTCTCTGTGATGTTGATGGTTTGGTATAGTAAACATGAGATATGATTAAAGGTGATTCAGGGATAG  
ATCAAGTGTCTGCCTAAGTAAATCTGGGNTTTCATTTTTTTTCTAGACATCTTGGA



**Claims**

1. An isolated polynucleotide which comprises a nucleotide sequence which has at least 70% identity to that of SEQ ID NOs:1 - 6 over the entire length of SEQ ID NOs:1 - 6; or  
5 a nucleotide sequence complementary to said isolated polynucleotide; and fragments of the polynucleotides.
2. The isolated polynucleotide as defined in claim 1 in which the identity is at least 95% to SEQ ID Nos: 1 - 6.
- 10 3. An isolated polynucleotide selected from:  
(a) the polynucleotide of any one of SEQ ID NOs:1- 6; and  
(b) a polynucleotide obtainable by screening an appropriate library under stringent hybridization conditions with a labeled probe having the sequence of any one of SEQ ID  
15 NOs:1 - 6 or a fragment thereof; and  
(c) a fragment of (a) or (b);  
or a nucleotide sequence complementary to said isolated polynucleotide
4. A process for diagnosing a disease or a susceptibility to a disease in a subject related  
20 to presence or expression or activity of a polynucleotide of any one of claims 1 to 3 in a subject comprising analyzing for the presence or amount of said polynucleotide in a sample derived from said subject.
5. A process for diagnosing the presence of colon cancer or a susceptibility to colon  
25 cancer in a subject related to presence or expression or activity of a polynucleotide of any one of claims 1 to 3 in a subject comprising analyzing for the presence or amount of said polynucleotide in a sample derived from said subject.

1/1

**Figure 1**  
**Hybridization blot for differential screening**

